

Sampled and Continuous-Time 1-Bit Signal Processing in CMOS for Wireless Sensor Networks

PhD Thesis

Håkon A. Hjortland
Department of Informatics
University of Oslo

January 2016



© Håkon A. Hjortland, 2016

*Series of dissertations submitted to the
Faculty of Mathematics and Natural Sciences, University of Oslo
No. 1721*

ISSN 1501-7710

All rights reserved. No part of this publication may be
reproduced or transmitted, in any form or by any means, without permission.

Cover on printed media: Hanne Baadsgaard Utigard.
Print production: Reprosentralen, University of Oslo.

*One or more included papers are copyright (C) IEEE.
One or more included papers are copyright (C) IET.
One or more included papers are copyright (C) Springer Science+Business Media.*

Preface

This dissertation has been submitted to Department of Informatics, Faculty of Mathematics and Natural Sciences, University of Oslo in partial fulfillment of the requirements for the degree of Doctor of Philosophy (PhD).

Acknowledgments

I would like to give great thanks to all my colleagues through the years of PhD work, who I have had interesting discussions and collaborations with: My supervisor Tor Sverre Lande and my co-supervisor Dag T. Wisland. The master students I co-supervised, Trygve K. Halvorsen, Øyvind Dahl, Olav E. Liseth, Daniel Mo, and Mats Risopatron Knutsen. My fellow PhD students Malihe Zarre Dooghabadi, Shanthi Sudalaiyandi, Tuan Anh Vu, Øystein Bjørndal, and post-doc Øyvind Næss. Technical responsible at the research group, Olav Stanly Kyrvestad. My colleagues at Novelda AS, Kjetil Meisal, Claus Limbodal, Olav E. Liseth, Kristian Granhaug, Nikolaj Andersen, Jørgen Andreas Michaelsen, and Mats Risopatron Knutsen. And if I have forgotten anyone, great thanks to you too. I would also like to thank everyone else that I have had discussions and collaborations with. Lastly, I would like to thank the NANO research group, Department of Informatics, Faculty of Mathematics and Natural Sciences, and University of Oslo.

During my PhD work, I was financed by Department of Informatics and also self-financed for a few years.

Håkon A. Hjortland
Oslo, January 2016

Abstract

This thesis investigates theoretical aspects of 1-bit signal processing and contains published papers describing many different CMOS implementations that utilize such processing. So-called Suprathreshold Stochastic Resonance (SSR) systems are studied in detail. These systems take a noisy signal as input and then 1-bit quantize it. As it turns out, the combination of noise, coarse quantization, and also integration, results in systems that, perhaps a bit unexpectedly, have quite good performance. This is known as the Stochastic Resonance (SR) phenomenon. Many interesting plots explaining different aspects of the behavior of SSR systems are shown. Detailed explanations of when SSR systems should and should not be used in, are given. This includes analysis of their power consumption. One of the most interesting results in this thesis, is that when the input to the SSR system has flicker noise, the SNR loss in the system will be negligible. With white Gaussian noise, the SNR loss will be much larger, namely 1.96 dB. The flicker-noise result does not seem to have been published before. In addition to using 1-bit signal processing, most of the CMOS implementations also utilize Impulse-Radio Ultra-WideBand (IR-UWB), and yet another commonality is that they can be used for Wireless Sensor Network (WSN) applications. Many of the implementations base much of their signal processing on Continuous-Time Binary-Value (CTBV) signals, which is an unusual way of doing things and therefore an interesting experiment. Some of the CMOS implementations are as follows: Radar, localization, communication, beamforming, bitstream processing, n-path band-pass filter, and power harvesting.

List of Publications

The following are chronological lists of the five journal articles and 21 conference papers that are presented in this thesis. See the table of contents or appendix A for a list grouped by topic.

Journal Articles

H. A. Hjortland and T. S. Lande. “CTBV Integrated Impulse Radio Design for Biomedical Applications”. *Biomedical Circuits and Systems, IEEE Transactions on*, Vol. 3, No. 2, pp. 79–88, Apr. 2009. doi:10.1109/TBCAS.2009.2014962.

O. E. Liseth, D. Mo, H. A. Hjortland, T. S. Lande, and D. T. Wisland. “Power-Efficient Cross-Correlation Beat Detection in Electrocardiogram Analysis Using Bitstreams”. *Biomedical Circuits and Systems, IEEE Transactions on*, Vol. 4, No. 6, pp. 419–425, Dec. 2010. doi:10.1109/TBCAS.2010.2079933.

M. Z. Dooghabadi, H. A. Hjortland, and T. S. Lande. “High precision calibrated digital delay element”. *Electronics Letters*, Vol. 47, No. 9, pp. 564–565, Apr. 28, 2011. doi:10.1049/el.2011.0395.

S. Sudalaiyandi, H. A. Hjortland, and T. S. Lande. “A continuous-time IR-UWB RAKE receiver for coherent symbol detection”. *Analog Integrated Circuits and Signal Processing*, Vol. 77, No. 1, pp. 17–27, Oct. 2013. doi:10.1007/s10470-013-0120-0.

M. Z. Dooghabadi, H. A. Hjortland, O. Nass, K. K. Lee, and T. S. Lande. “An IR-UWB Transmitter for Ranging Systems”. *Circuits and Systems II: Express Briefs, IEEE Transactions on*, Vol. 60, No. 11, pp. 721–725, Nov. 2013. doi:10.1109/TCSII.2013.2281909.

Conference Papers

H. A. Hjortland, D. T. Wisland, T. S. Lande, C. Limbodal, and K. Meisal. “CMOS Impulse Radar”. In: *Norchip Conference, 2006. 24th*, pp. 75–79, Linköping, Nov. 2006. doi:10.1109/NORCHP.2006.329248.

H. A. Hjortland, D. T. Wisland, T. S. Lande, C. Limbodal, and K. Meisal. “Thresholded samplers for UWB impulse radar”. In: *Circuits and Systems, 2007. ISCAS 2007. IEEE International Symposium on*, pp. 1210–1213, New Orleans, LA, May 2007. doi:10.1109/ISCAS.2007.378326.

T. S. Lande and H. A. Hjortland. “Impulse Radio technology for Biomedical applications”. In: *Biomedical Circuits and Systems Conference, 2007. BIOCAS 2007. IEEE*, pp. 67–70, Montreal, Que., Nov. 2007. doi:10.1109/BIOCAS.2007.4463310.

T. K. Halvorsen, H. A. Hjortland, and T. S. Lande. “Power Harvesting Circuits in 90 nm CMOS”. In: *NORCHIP, 2008.*, pp. 154–157, Tallinn, Nov. 2008. doi:10.1109/NORCHIP.2008.4738301.

O. Dahl, H. A. Hjortland, T. S. Lande, and D. T. Wisland. “Close range impulse radio beamformers”. In: *Ultra-Wideband, 2009. ICUWB 2009. IEEE International Conference on*, pp. 205–209, Vancouver, BC, Sep. 2009. doi:10.1109/ICUWB.2009.5288755.

O. E. Liseth, H. A. Hjortland, and T. S. Lande. “Power efficient cross-correlation beat detection in electrocardiogram analysis using bitstreams”. In: *Biomedical Circuits and Systems Conference, 2009. BioCAS 2009. IEEE*, pp. 237–240, Beijing, Nov. 2009. doi:10.1109/BIOCAS.2009.5372041.

O. E. Liseth, D. Mo, H. A. Hjortland, T. S. Lande, and D. T. Wisland. “Power efficient cross-correlation using bitstreams”. In: *NORCHIP, 2009*, pp. 1–4, Trondheim, Nov. 2009. doi:10.1109/NORCHIP.2009.5397844.

M. Z. Dooghabadi, T. A. Vu, S. Sudalaiyandi, H. A. Hjortland, T. S. Lande, O. Naess, and S. E. Hamran. “Electromagnetic impulse radio camera”. In: *NORCHIP, 2009*, pp. 1–5, Trondheim, Nov. 2009. doi:10.1109/NORCHIP.2009.5397835.

T. A. Vu, S. Sudalaiyandi, M. Z. Dooghabadi, H. A. Hjortland, O. Nass, T. S. Lande, and S. E. Hamran. “Continuous-time CMOS quantizer for ultra-wideband applications”. In: *Circuits and Systems (ISCAS), Proceedings of 2010 IEEE International Symposium on*, pp. 3757–3760, Paris, France, May 2010. doi:10.1109/ISCAS.2010.5537745.

S. Sudalaiyandi, M. Z. Dooghabadi, T. A. Vu, H. A. Hjortland, Ø. Nass, T. S. Lande, and S. E. Hamran. “Power-efficient CTBV symbol detector for UWB applications”. In: *Ultra-Wideband (ICUWB), 2010 IEEE International Conference on*, pp. 1–4, Nanjing, Sep. 2010. doi:10.1109/ICUWB.2010.5614422.

K. K. Lee, M. Z. Dooghabadi, H. A. Hjortland, O. Naess, and T. S. Lande. “A novel 6.5 pJ/pulse impulse radio pulse generator for RFID tags”. In: *Circuits and Systems (APCCAS), 2010 IEEE Asia Pacific Conference on*, pp. 184–187, Kuala Lumpur, Dec. 2010. doi:10.1109/APCCAS.2010.5775001.

K. K. Lee, M. Z. Dooghabadi, H. A. Hjortland, O. Næss, and T. S. Lande. “A 5.2 pJ/pulse impulse radio pulse generator in 90 nm CMOS”. In: *Circuits and Systems (ISCAS), 2011 IEEE International Symposium on*, pp. 1299–1302, Rio de Janeiro, May 2011. doi:10.1109/ISCAS.2011.5937809.

T. A. Vu, S. Sudalaiyandi, H. A. Hjortland, O. Naess, T. S. Lande, and S. E. Hamran. “A 70-dB, 3.1-10.6-GHz CMOS amplifier in low-power 90 nm CMOS”. In: *Circuits and Systems (MWSCAS), 2011 IEEE 54th International Midwest Symposium on*, pp. 1–4, Seoul, Aug. 2011. doi:10.1109/MWSCAS.2011.6026639.

D. T. Wisland, S. Stoa, N. Andersen, K. Granhaug, T.-S. Lande, and H. A. Hjortland. “CMOS nanoscale impulse radar utilized in 2-dimensional ISAR imaging system”. In: *Radar Conference (RADAR), 2012 IEEE*, pp. 0714–0719, Atlanta, GA, May 2012. doi:10.1109/RADAR.2012.6212231.

S. Sudalaiyandi, T.-A. Vu, H. A. Hjortland, O. Nass, and T.-S. Lande. “Continuous-time single-symbol IR-UWB symbol detection”. In: *SOC Conference (SOCC), 2012 IEEE International*, pp. 198–201, Niagara Falls, NY, Sep. 2012. doi:10.1109/SOCC.2012.6398347.

T. A. Vu, S. Sudalaiyandi, H. A. Hjortland, O. Nass, T. S. Lande, and S. E. Hamran. “An inductorless 6-Path band-pass filter with tunable center frequency for UWB applications”. In: *Ultra-Wideband (ICUWB), 2012 IEEE International Conference on*, pp. 164–167, Syracuse, NY, Sep. 2012. doi:10.1109/ICUWB.2012.6340442.

S. Sudalaiyandi, H. A. Hjortland, T. Vu, O. Næss, and T. S. Lande. “Continuous-time high-precision IR-UWB ranging-system in 90 nm CMOS”. In: *Solid State Circuits Conference (A-SSCC), 2012 IEEE Asian*, pp. 349–352, Kobe, Nov. 2012. doi:10.1109/ASSCC.2012.6570786.

S. Sudalaiyandi, H. A. Hjortland, T.-A. Vu, O. Naess, and T. S. Lande. “Live demonstration: IR-UWB transceiver with localization based on ToF technique suitable for wireless capsule endoscopy”. In: *Biomedical Circuits and Systems Conference (BioCAS), 2012 IEEE*, pp. 89–89, Hsinchu, Nov. 2012. doi:10.1109/BioCAS.2012.6418495.

S. Sudalaiyandi, T.-A. Vu, H. A. Hjortland, O. Naess, and T. S. Lande. “Continuous-time symbol detector for IR-UWB rake receiver in 90 nm CMOS”. In: *Circuits and Systems (APCCAS), 2012 IEEE Asia Pacific Conference on*, pp. 487–490, Kaohsiung, Dec. 2012. doi:10.1109/APCCAS.2012.6419078.

M. Z. Dooghabadi, H. A. Hjortland, and T. S. B. Lande. “An ultra-wideband receiving antenna array”. In: *Circuits and Systems (ISCAS), 2013 IEEE International Symposium on*, pp. 2373–2376, Beijing, China, May 2013. doi:10.1109/ISCAS.2013.6572355.

M. Z. Dooghabadi, H. A. Hjortland, and T. S. B. Lande. “A linear IR-UWB MIMO radar array”. In: *Ultra-Wideband (ICUWB), 2013 IEEE International Conference on*, pp. 13–19, Sydney, NSW, Sep. 2013. doi:10.1109/ICUWB.2013.6663814.

Contents

Preface	3
Acknowledgments	3
Abstract	5
List of Publications	7
Contents	11
Introduction	15
Motivation	15
Aims of This Study	15
Research Questions	15
Research Methods	15
Typographical Choices	15
Outline of Thesis	16
1 Wireless Sensor Networks	17
1.1 Introduction	18
2 Ultra-Wideband	21
2.1 Introduction	22
2.1.1 Health Effects of RF Waves	23
3 1-Bit Signal Processing	25
3.1 Introduction	27
3.2 Central Concepts and Usage Classes	30
3.3 Representing Analog Signals	32
3.3.1 Signal Conversion	32
3.3.2 Signal Processing Operations	35
3.3.3 Overview of Basic Modulation Types, Converters, and Operations . . .	35
3.4 Power Consumption	39
3.4.1 Unit-Element Amplifier	40
3.4.2 Unit-Element Amplifier in Sampled Systems	46
3.4.3 Analog, Multi-Bit, and 1-Bit Modulations/Encodings	54

3.4.4	Power Benefits of Digital Systems	60
3.4.5	Power Benefits of 1-Bit Systems	60
3.5	Example Applications	61
3.6	1-Bit Signal Processing As a Fundamental Principle	62
4	Sampled 1-Bit Signal Processing	63
4.1	Introduction	65
4.2	Other Uses for 1-Bit Sequences Than Conveying Signals	66
4.3	Suprathreshold Stochastic Resonance	67
4.3.1	Stochastic Resonance	67
4.3.2	Subthreshold Stochastic Resonance	69
4.3.3	Suprathreshold Stochastic Resonance	69
4.3.4	Trading off SNR and Sampling Rate	73
4.3.5	Effects of Noise, Nonlinearity, and Integration	73
4.4	Comparison with Sigma-Delta Modulation	77
4.5	Threshold Levels	78
4.5.1	Swept-Threshold Sampling	78
4.6	Integration by Oversampling or by Parallelization	80
4.7	General Benefits of Oversampling	81
4.7.1	Oversampling in the Retina	82
4.8	Decoding	84
4.9	SNR Metrics	85
4.9.1	Measurement Setup	85
4.9.2	Gains	86
4.9.3	SNR Metric A: Consider Error in Absolute Value	88
4.9.4	SNR Metric B: Maximize SNR Value	89
4.9.5	SNR Metric C: Correctly Identify Signal Part	91
4.9.6	Examples of the Differences	92
4.9.7	Oversampling	92
4.9.8	Distortion Compensation	93
4.9.9	Limitations of the SNR Metric	93
4.10	Small-Signal Analysis	94
4.11	Large-Signal Analysis	100
4.11.1	With an Input Signal Distribution	100
4.11.2	Integration	101
4.11.3	Uniform Noise	101
4.12	SNR As a Function of Input Noise	105
4.12.1	SNR Metrics A, B, C	106
4.12.2	Quantizer Output Levels	106
4.12.3	Related Work	110
4.13	Colored Noise	112
4.13.1	Why SSR with Colored Noise Works	113
4.13.2	White Noise from First-Order Anti-Aliasing Filter	114

4.14	Spectrum Simulations	115
4.14.1	Measuring the SNR	115
4.14.2	Overview	117
4.14.3	Simulations	117
4.15	Bitstream Processing	134
4.16	Example Applications	135
5	Continuous-Time 1-Bit Signal Processing	137
5.1	Introduction	138
5.1.1	Example Applications	138
5.2	Zierhofer Modulation	140
5.3	CTBV-Based Logic	141
6	Circuit Solutions	143
6.1	Introduction	144
7	Conclusions	147
7.1	Conclusions	148
7.2	Main Contributions of This Work	148
7.2.1	General Overview of 1-Bit Signal Processing	148
7.2.2	Classification of 1-Bit Concepts	148
7.2.3	Energy Consumption per Sample	148
7.2.4	Power Comparison of Modulations/Encodings	149
7.2.5	Forced Oversampling in Analog Systems	149
7.2.6	1-Bit Signal Processing As a Fundamental Principle	149
7.2.7	SNR Metrics	149
7.2.8	Extra Strong SR Peak	149
7.2.9	Large-Signal Analysis	149
7.2.10	SNR As a Function of Input Noise	149
7.2.11	Why SSR with Colored Noise Works	150
7.2.12	White Noise from First-Order Anti-Aliasing Filter	150
7.2.13	Spectrum Simulations	150
7.2.14	Practical SSR System	150
7.2.15	Flicker Noise Gives 0 dB SNR Loss	150
7.2.16	Continuous-Time SSR Without Noise	150
7.2.17	SSR in Evolution	151
7.2.18	CTBV Symbol Detector	151
7.2.19	High-Precision Timing Calibration	151
7.3	Future Work	151
A	Papers	153
A.1	Radar and CTBV	155
Paper 1	CMOS Impulse Radar	157
Paper 2	Thresholded samplers for UWB impulse radar	165

Paper 3	Impulse Radio technology for Biomedical applications	171
Paper 4	CTBV Integrated Impulse Radio Design for Biomedical Applications	177
Paper 5	CMOS nanoscale impulse radar utilized in 2-dimensional ISAR imaging system	189
A.2	Localization and Communication	197
Paper 6	Power-efficient CTBV symbol detector for UWB applications	199
Paper 7	Continuous-time single-symbol IR-UWB symbol detection	205
Paper 8	Continuous-time high-precision IR-UWB ranging-system in 90 nm CMOS	211
Paper 9	Live demonstration: IR-UWB transceiver with localization based on ToF technique suitable for wireless capsule endoscopy	217
Paper 10	Continuous-time symbol detector for IR-UWB rake receiver in 90 nm CMOS	221
Paper 11	A continuous-time IR-UWB RAKE receiver for coherent symbol detection	227
Paper 12	An IR-UWB Transmitter for Ranging Systems	241
A.3	Beamforming	249
Paper 13	Close range impulse radio beamformers	251
Paper 14	Electromagnetic impulse radio camera	259
Paper 15	High precision calibrated digital delay element	267
Paper 16	An ultra-wideband receiving antenna array	271
Paper 17	A linear IR-UWB MIMO radar array	277
A.4	Bitstream Processing	287
Paper 18	Power efficient cross-correlation beat detection in electrocardiogram analysis using bitstreams	289
Paper 19	Power efficient cross-correlation using bitstreams	295
Paper 20	Power-Efficient Cross-Correlation Beat Detection in Electrocardiogram Analysis Using Bitstreams	301
A.5	Pulse Generator, N-Path Filter, Quantizer, and Power Harvesting	311
Paper 21	A novel 6.5 pJ/pulse impulse radio pulse generator for RFID tags . .	313
Paper 22	A 5.2 pJ/pulse impulse radio pulse generator in 90 nm CMOS	319
Paper 23	An inductorless 6-Path band-pass filter with tunable center frequency for UWB applications	325
Paper 24	Continuous-time CMOS quantizer for ultra-wideband applications .	331
Paper 25	A 70-dB, 3.1-10.6-GHz CMOS amplifier in low-power 90 nm CMOS	337
Paper 26	Power Harvesting Circuits in 90 nm CMOS	343

Acronyms	349
-----------------	------------

Bibliography	353
---------------------	------------

Introduction

This thesis investigates 1-bit signal processing, from theory to feasible implementations. It will be used in technology for Wireless Sensor Network (WSN) which will be implemented in CMOS. IR-UWB will also be utilized.

Motivation

We want to understand the theoretical aspects of 1-bit signal processing better, so that we may use this e.g. in CMOS hardware.

We want to explore the CTBV paradigm and 1-bit signal processing in CMOS for IR-UWB. Can this bring something new to the field? We also want to implement other technologies for WSN nodes.

Aims of This Study

To produce new theoretical knowledge about 1-bit signal processing.

To create many novel circuit solutions, hoping some of them will turn out to be of academic or practical interest or that some interesting observations can be made.

Research Questions

CTBV is not very much in use, so it warrants research in the form of exploration, analysis, and gathering experience about practical issues.

Research Methods

Make analytical derivations, numerical calculations, and simulations about 1-bit signal processing.

Come up with ideas for circuit solutions and test them in CMOS.

Typographical Choices

References to the included papers are shown like this: Paper 4. Wikipedia links are shown like this: Dynamical system [W]. Citations include DOI links or web links if available: [Hjor 09 ©]

[Hjor 06 ©].

Outline of Thesis

Chapter 1 and 2 briefly introduce WSN and Ultra-WideBand (UWB). Chapter 3 gives an overview of 1-bit signal processing, and chapter 4 looks into the details of sampled 1-bit systems. Chapter 5 briefly discusses continuous-time 1-bit systems. Chapter 6 gives a quick overview of the circuit solutions in the published papers. Chapter 7 concludes the thesis. Appendix A contains the published papers.

Chapter 1

Wireless Sensor Networks

Contents

1.1	Introduction	18
-----	------------------------	----

1.1 Introduction

Wireless Sensor Networks (WSNs) [Akyi 02 ☉, Stan 08 ☉] consist of anywhere from a few to thousands of simple sensor nodes spread over an area that gather data, perhaps process it, and then transmit it over a wireless link. The sensor nodes are also known as motes, and they can be designed to form self-organizing networks, although simpler network structures are possible too. The networks can be heterogeneous, comprising elements such as static and mobile sensor nodes, user interface nodes, actuator nodes, gateway nodes, Internet infrastructure, centralized control systems, robots to move and charge nodes, and so forth. The sensor nodes should be small, cheap, robust, independent, and have a long battery life or use energy harvesting. Different applications will require different properties such as high data rate, autonomous decision making, secure data exchange, robust communication, low power consumption, or precise position information.

There is a wide array of possible applications for WSNs [Aram 05 ☉]. They can be used by the military to monitor enemy movement. They can be used to monitor the environment indoors or outdoors, such as detecting fires, monitoring the environment and wildlife in nature, and providing useful data about crops in agriculture. They can also be used to monitor electricity usage in homes, i.e. be a part of the smart grid [W]. They can be used for logistics, such as for tracking items. They can be used for monitoring e.g. temperature and vibration in machinery. They can be used in the form of wearable motes by firemen to track their position and condition during dangerous operations. They can be used in health care, such as in assisted-living facilities for senior citizens where they can be used for fall detection, to monitor medication intake, and to detect diseases from behavior patterns. In general, WSNs might be useful in any application where there is interesting information to be found over a large area or volume and where remote sensing such as cameras can not do the job.

There are several other concepts that are more or less related to WSNs. One of these is Internet of Things (IoT) [W] and the related concepts smart grids, smart homes, intelligent transportation, and smart cities. Another is Machine to Machine (M2M) [W]. Finally, WSNs are also a type of ubiquitous computing [W], also known as pervasive computing, ambient intelligence, ambient media, “everyware”, physical computing, haptic computing, and “things that think”.

Some of the key elements a WSN sensor node might consist of are as follows and as also

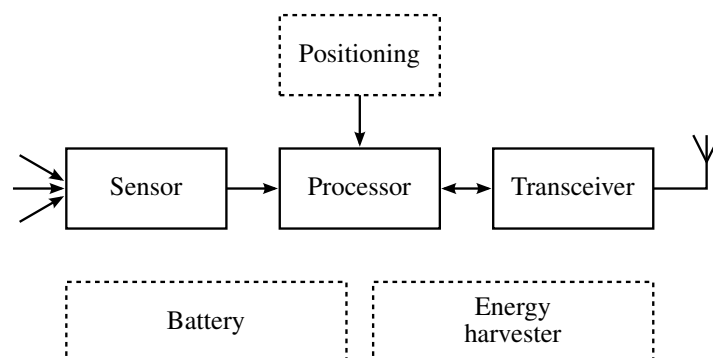


Figure 1.1: A Wireless Sensor Network (WSN) sensor node.

shown in figure 1.1:

- Power harvesting
- Sensors
- Processing
- Communication
- Positioning

These elements and the sensor node as a whole might have different desired functionality and properties depending on the application. It might be desirable that the sensor node is e.g. small, cheap, uses low power, has high-precision positioning, pre-processes the data before transmission, combines positioning/communication/sensing in one antenna, has robust communication, or that it has higher-level functions such as mesh network routing protocols. Each of these points might give strong indications about what kind of technologies and solutions might work and not, and they also each represent interesting research challenges.

UWB, CMOS, and 1-bit signal processing work well together and can be used to solve many of the aforementioned challenges. WSNs therefore present themselves as suitable practical applications for these three interesting fields of research. Because of this, we have chosen WSNs as an umbrella motivation for the research we have done in this thesis into these technologies. We have, in varying degrees, implemented in silicon each of the elements in the bullet-point list above. This is presented in the papers in appendix A. Additionally, a theoretical treatment of 1-bit signal processing is presented in the chapters of this thesis. Note that we have focused on sensor node hardware and not higher-level issues such as mesh network routing protocols in this thesis.

Chapter 2

Ultra-Wideband

Contents

2.1	Introduction	22
2.1.1	Health Effects of RF Waves	23

2.1 Introduction

Ultra-WideBand (UWB) [W] is a radio technology with very high bandwidth. This concept enables many new and interesting applications. As an example, the Federal Communications Commission (FCC) regulations allow a transmit power of -41.3 dBm/MHz in the frequency band from 3.1 GHz to 10.6 GHz. This gives a total bandwidth of 7.5 GHz and a total power of about 0.5 mW. Other countries have different spectrum regulations [Fern 10 ♻]. The basic idea behind allowing such wideband transmission, is that the power received by any given narrowband channel will be so low that it will not be a problem.

Since the transmit power is so low, a UWB system will only have a quite short operating range. This means that UWB has a high spatial capacity [W], i.e. that many users can each utilize a high data rate without needing to be too far away from each other. A system with higher transmit power would cause interference over a wider area, and it would therefore allow fewer users in a given area and would consequently have a lower spatial capacity. Additionally, UWB systems do not need to use high amplitudes in order to achieve high data rates, as in e.g. 256-QAM systems, since UWB systems can achieve high data rates through the use of the wide bandwidth. UWB systems can therefore have a low SNR. This means that such systems do not need to be far away from each other before the MSB of an interfering system becomes weaker than the LSB of the intended signal. A 256-QAM system, on the other hand, has an MSB which is much stronger than the LSB, and therefore needs to be far away from other users in order to prevent interference. UWB systems therefore get a higher spatial capacity than narrowband systems not only due to their low power but also due to this particular effect.

Due to the large bandwidth, the UWB signals will have a very high time resolution. This makes UWB suitable for radar and localization. UWB radars can be realized by transmitting short pulses. Such impulse radars do not need any pulse compression techniques, which might help reduce the system complexity. UWB radars can be used for a wide range of sensing applications. The low power, the material-penetration properties of RF signals, and the high resolution make such systems interesting competitors to other sensing modalities and narrowband radar systems. The low-power non-ionizing radiation makes such systems suitable for medical applications too. Section 2.1.1 below discusses the health effects of RF radiation in greater detail.

UWB systems are also suitable for communication. The large bandwidth is a good argument itself since this allows very high data rates. Another advantage is the channel diversity offered by the large bandwidth. To put it one way, destructive multipath interference at a few frequencies is not a problem since there are many other frequencies available. To put it another way, a UWB impulse in an impulse-based system that experiences multipath reflections basically just gets free retransmissions, which can in fact be exploited. Yet another advantage of UWB communication systems is that low SNR and high bandwidth is more energy efficient than high SNR and low bandwidth, as can be seen from the Shannon-Hartley channel capacity theorem [W], i.e. $C = B \cdot \log_2(1 + \text{SNR})$. Doubling the SNR power ratio will only add one more bit, whereas doubling the bandwidth will double the data rate. It is therefore more energy efficient to reduce the SNR and instead use a larger bandwidth.

In the papers in appendix A, we will present many different UWB solutions implemented in CMOS, such as radar, localization, communication, and beamforming.

2.1.1 Health Effects of RF Waves

According to a literature review by IEEE International Committee on Electromagnetic Safety, any adverse health effects from 3 kHz–300 GHz EM radiation is only due to excessive heating, not any other mechanisms such as ionization [Wu 15 ☉, p. 66]. So if the exposure level is low enough to avoid any dangerous rise of temperature in the body, the radiation should be safe. The FCC and World Health Organization (WHO) have made more precautionary statements.

One possible non-thermal effect is that low-megahertz radiation can reorient dipolar molecules, but this is not known to occur at higher frequencies [Wu 15 ☉, p. 81]. For millimeter waves, i.e. 30 GHz to 300 GHz EM radiation, there are reports of some effects on cell membranes, but also reports of some positive effects in cancer treatment, on the immune system, and for wound healing [Wu 15 ☉, pp. 79]. More research is required before anything conclusive can be said one way or another, though.

Since UWB systems generally have a transmit power of less than a milliwatt, there is good reason to believe that they will not have any detrimental health effects.

Chapter 3

1-Bit Signal Processing

Contents

3.1	Introduction	27
3.2	Central Concepts and Usage Classes	30
3.3	Representing Analog Signals	32
3.3.1	Signal Conversion	32
3.3.2	Signal Processing Operations	35
3.3.3	Overview of Basic Modulation Types, Converters, and Operations	35
3.4	Power Consumption	39
3.4.1	Unit-Element Amplifier	40
3.4.1.1	Power Consumption During Continuous Operation	41
3.4.1.2	Real Amplifiers Compared to the Ideal Amplifier	42
3.4.1.3	Three Distinct Energy Quantities	42
3.4.1.4	The Ideal Amplifier Gives Exact Energy per Degree of Freedom	43
3.4.1.5	Proof of $kT/2$	43
3.4.1.6	Remaining Questions About $kT/2$	45
3.4.2	Unit-Element Amplifier in Sampled Systems	46
3.4.2.1	Integration by Parallelization	46
3.4.2.2	Frequencies and Related Quantities	46
3.4.2.3	Signal, Noise, and Dynamic Range	47
3.4.2.4	Power Consumption During Continuous Operation	47
3.4.2.5	Energy Consumption per Sample	47
3.4.2.6	Energy Consumption of Inverters	48
3.4.2.7	Discussion of the Expression for Energy Consumption	49
3.4.2.8	Overall Power Consumption	50
3.4.2.9	The Effects of Tuning the Different Parameters	50
3.4.2.10	Calculated Results for CMOS Inverters	51
3.4.3	Analog, Multi-Bit, and 1-Bit Modulations/Encodings	54
3.4.3.1	Unit-Element Amplifier	54
3.4.3.2	High-SNR Signals	55
3.4.3.3	Low-SNR Signals	57
3.4.3.4	Forced Oversampling in Analog Systems	58
3.4.3.5	Summary	59
3.4.4	Power Benefits of Digital Systems	60
3.4.5	Power Benefits of 1-Bit Systems	60
3.5	Example Applications	61
3.6	1-Bit Signal Processing As a Fundamental Principle	62

3.1 Introduction

1-bit signal processing enables simple yet advanced systems. Despite the coarse quantization, the SNR loss can be very low when converting a signal to 1-bit precision. This signal processing paradigm is suitable for WSN nodes due to its simplicity and for UWB applications since they both have low SNR and high bandwidth. The current chapter gives an overview of 1-bit signal processing, and chapter 4 and 5 respectively detail the sampled case and the continuous-time case, in particular investigating the effects of 1-bit quantization of signals. Some variations of such 1-bit quantization are known as Suprathreshold Stochastic Resonance (SSR). The role of 1-bit signals in physics, biology, and electronics is explored too. Several CMOS system implementations utilizing Continuous-Time Binary-Value (CTBV) signal processing are presented in appendix A.

As the supply voltage of fine-pitch technologies decreases, analog signal processing becomes more and more difficult to perform due to the shrinking headroom. On the other hand, though, the circuits become faster and more plentiful, which opens up new opportunities for signal processing. To exploit this fact, we have to move away from the analog signal processing domain and on to a new paradigm which is able to utilize these properties. 1-bit signal processing is a perfect candidate for this. Its simpler circuits, high clock frequency, and parallelization makes it a good fit for fine-pitch technology. Moving from analog to 1-bit signal processing can give many benefits, starting with those we get simply by switching to digital signal processing:

- Might be implementable in a given technology, e.g. CMOS, when corresponding analog solutions might not be implementable
- Smaller area
- Room for parallelization
- Possibly lower power consumption
- Allows the creation of more advanced systems
- Can store and transfer the signal without loss over time
- Immunity to noise, PVT variations, and device mismatch
- Better linearity, but just within the digital domain. Linearity of A/D and D/A conversion still challenging.
- Processing capacity can be increased by using a clock frequency which is higher than the Nyquist rate or by creating parallel circuits

Furthermore, moving from multi-bit to 1-bit signal processing gives several benefits of its own:

- Simpler
- Faster
- Fewer wires
- Timing skew between the parallel bits of a bus is no longer an issue
- Smaller area
- Room for parallelization
- Possibly lower power consumption

- Integration and the use of a stochastic paradigm gives some immunity to noise, nonlinearity, device mismatch, bit errors, and faulty elements
- Both the ADC, the processing, and the DAC become simpler and smaller
- 1-bit ADCs and DACs do not have the same problem with nonlinearity as multi-bit versions do
- Well suited for integration of the results from many low-SNR sensors. The increased integration could make the end result better than the result from a single high-quality analog sensor.
- Continuous-time 1-bit systems provide high time resolution without the use of high-frequency clocks

There are also some potential drawbacks with 1-bit signal processing, though, which it is important to be aware of. Firstly, a 1-bit system might require a clock, but this does not necessarily need to be an external clock reference such as a crystal. A simple internal oscillator could suffice. Secondly, the system might require oversampling or parallelization in order to compensate for the coarse quantization. This incurs a penalty in the form of increased power consumption or area usage. Thirdly, some operations, such as compression, clipping, filtering, and summing, could be more challenging to perform in a 1-bit system than in an analog system. Finally, SSR systems, which are the focus of chapter 4, only handle input with low or moderate SNR, and they also cause some SNR loss.

How can we turn an analog system into a 1-bit system? In particular, how do we turn it into a simple SSR system? SSR systems require an SNR per sample, i.e. a full-band SNR, of less than about 0 dB in order to work well. The in-band SNR can still be above 0 dB, though. This will be explained in chapter 4. To transform an analog system into an SSR system, we therefore have to reduce the SNR per sample. One way of doing this is to split the large transistors of the LNA in an analog system into several smaller transistors. This gives us many smaller and noisier input amplifiers, each with an SNR per sample which is below 0 dB. The output from each amplifier can then be 1-bit quantized, which will actually not cause much SNR loss, as will also be explained in chapter 4. We then have an SSR system with integration by parallelization. A different way to transform an analog system is to use oversampling. The higher sampling rate will cause more high-frequency thermal noise to be sampled, which will reduce the SNR per sample to less than 0 dB while leaving the in-band SNR unchanged. The samples are then 1-bit quantized. This gives us an SSR system with integration by oversampling. Note how the input amplifier stage uses the same amount of power for both the analog system and the two 1-bit systems.

Here we notice an essential point about 1-bit systems, namely that we compensate for the coarse value quantization by increasing the temporal resolution, i.e. by oversampling, or by using multiple signal channels, i.e. by utilizing parallelization. In other words, we find some way to do integration in order to achieve an SNR which is higher than the quite low native SNR provided by 1-bit signals without integration.

It is tempting to think that we can improve the SNR of a 1-bit system beyond that of an analog system by increasing the amount of integration. However, if we do integration by parallelization, more integration would mean more input amplifiers and thus a larger capacitive load on the input node and a higher power consumption. An analog system could do just this too, so the 1-bit

system does not provide anything special in this regard. If we do integration by oversampling, though, we note that an increase in sampling rate will not cause any increase in the in-band SNR, so this does not help either. There is one case where we can improve the SNR by integrating more, though, and that is if we have many parallel sensors. Each sensor could have an in-band SNR of slightly less than 0 dB, and we could then sample each of them at the Nyquist rate and 1-bit quantize the samples. Each sensor and input stage could be very small and use very little power with such a 1-bit architecture, so we could have very many of them without using too much area. It is important to note that the power consumption will scale with the end result SNR, though. A multi-sensor system like this, under the name Polarity Coincidence Array (PCA), has been proposed for sonar applications [Kane 66 ⑤].

3.2 Central Concepts and Usage Classes

There are many different concepts relating to 1-bit signal processing, such as bitstreams, Pulse-Density Modulation (PDM), and SSR. As it turns out, many of these concepts can be classified into three classes, namely underlying data carriers, basic modulation types / basic data formats, and modulators / signal converters. Such a classification gives a clearer understanding of how the different concepts relate to each other. Figure 3.1 illustrates the classification and lists several 1-bit concepts, also showing how they relate.

Bitstreams and CTBV signals are underlying data carriers, i.e. they are the signaling capacity provided by the underlying hardware which the rest of the system design is built upon. The distinction between sampled and continuous-time 1-bit data carriers is important, and the two will be discussed separately in chapter 4 and 5 respectively.

PDM and Pulse-Position Modulation (PPM) are basic modulation types, i.e. they use an underlying data carrier to represent a signal by means of a modulation. There are many ways to create e.g. a PDM signal, and that is why we use this term to stress that basic modulation types are not necessarily actual modulation techniques, they are just a common format which could be shared by several such techniques.

SSR, Sigma-Delta Modulation (SDM), and Pulse-Width Modulation (PWM) are modulators, i.e. they convert input data to a basic modulation type. Some of them might convert both from the same data type and to the same basic modulation type, but if so, they will each have different properties such as what the bitstream sequence looks like. They will therefore be suitable for different purposes.

A different classification which is also important, are the usage classes. Several of these can be seen in figure 3.1. An important usage class for 1-bit signal processing, both sampled and continuous-time, is to work with representations of analog signals, i.e. to perform A/D and D/A conversions and to process signals in the 1-bit domain. This will be discussed in the next section and is a central theme in the thesis.

A different usage class, which uses sampled 1-bit signal processing, i.e. bitstreams, is serial links such as Serial Peripheral Interface (SPI), Universal Serial Bus (USB), Serial ATA (SATA), and Peripheral Component Interconnect Express (PCIe). 1-bit signals are suitable for high-speed buses like these because parallel buses suffer from timing skew between the bits, which will limit the achievable speeds. A serial link will obviously not have this problem, and by embedding the clock signal with the data bits, timing skew between clock and data bits will not be an issue either. Additionally, a 1-bit bus needs fewer wires, which reduces cost and cable size. Because of these advantages, a 1-bit architecture is often chosen for wired data links. Another usage class for bitstreams, is 1-bit CPUs, a technique which reduces area consumption at the cost of requiring a greater number of clock cycles to execute each instruction.

A usage class for CTBV signals, is clock, event, trigger, and timing signals. We could use the rising edge of a continuous-time 1-bit signal for this. Sometimes the duration, i.e. the pulse width, is of interest too. Since no time discretization happens, the timing precision will be very good. A different usage class for CTBV signals, is to threshold analog signals with a high SNR. This could be used to detect events or to sweep through a repeated signal in order to reconstruct it, as is done with the Swept Threshold (ST) technique, which will be described in section 4.5.1.

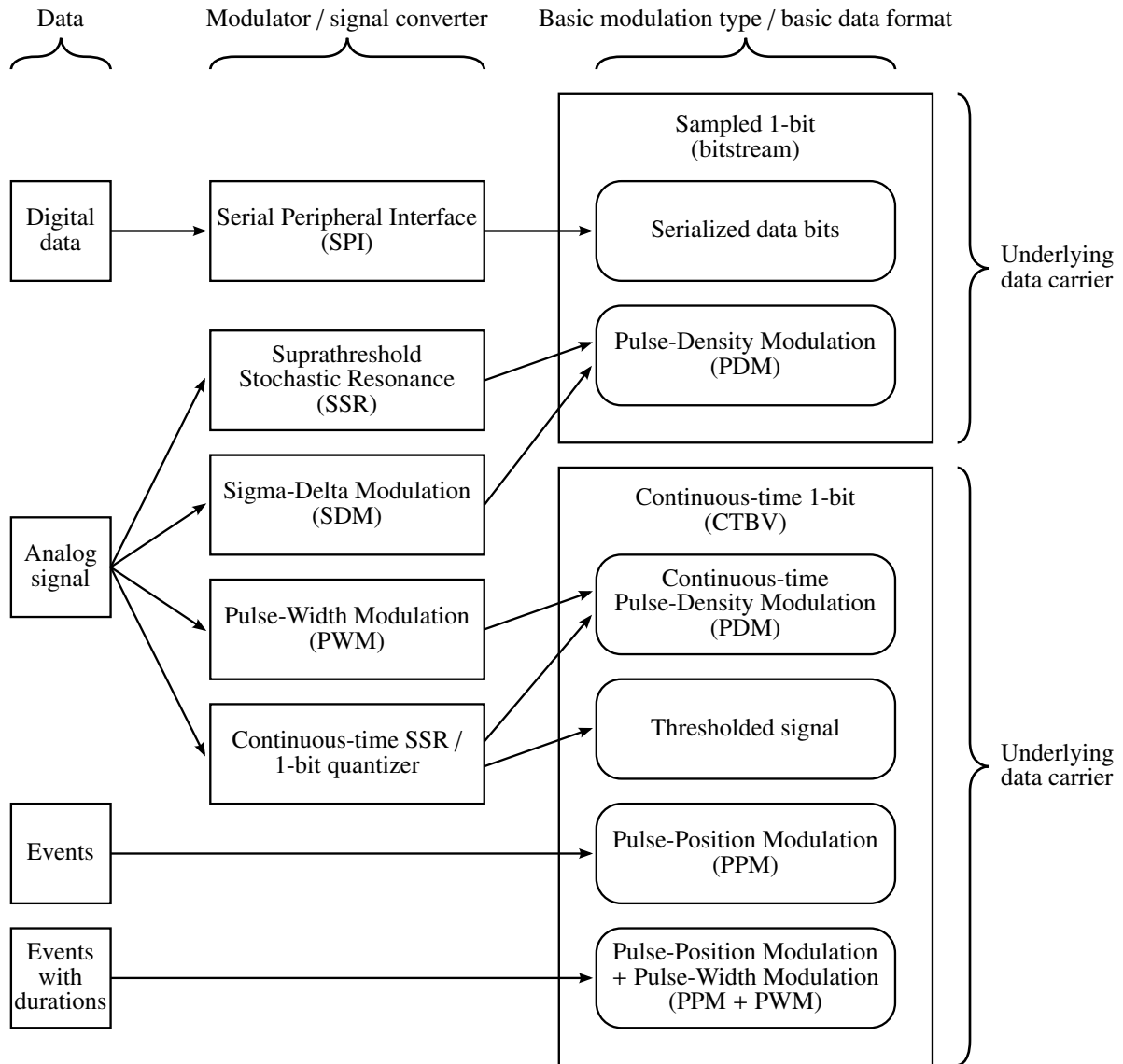


Figure 3.1: Overview of how some various 1-bit concepts relate to each other and how they can be classified as underlying data carriers, basic modulation types, and modulators.

3.3 Representing Analog Signals

1-bit signals can be used to represent analog signals. This means that we can somehow convert an analog signal to a 1-bit signal and that we can then somehow convert it to an analog signal again, i.e. there exist A/D and D/A converters that are able to transform a signal between the domains without causing too much loss. In addition to this basic conversion back and forth, we are also able to perform some signal processing operations on the analog signal while it is still in its 1-bit form. The 1-bit signal processing paradigm is therefore interesting both for ADC, DAC, and processing applications. Note that we could use a multi-bit signal just as well as an actual analog signal.

3.3.1 Signal Conversion

One interesting property of 1-bit ADCs and DACs is that they are inherently linear. Multi-bit converters have many threshold levels and output levels, which means that they have to be designed carefully to make sure that the levels are evenly spread, i.e. that the converter has a linear response. A 1-bit converter, on the other hand, has either one threshold level or two output levels, which makes it impossible for them to have unevenly spread levels. 1-bit circuits are also generally simpler than multi-bit circuits. These properties contribute to the popularity of sigma-delta converters.

There are many different modulations which can be used to convert an analog signal to a 1-bit signal. In this thesis, we will focus on modulations which are types of Pulse-Density Modulation (PDM). PDM means that in the spectrum of such a pulse train signal, the signal of interest is located at lower frequencies, possibly with some noise both inside and outside the signal band. Pulse-Frequency Modulation (PFM) with 50 % duty cycle and Pulse-Position Modulation (PPM) are examples of modulations that could represent an analog signal but which are not types of PDM. The three 1-bit signal modulations Suprathreshold Stochastic Resonance (SSR), Sigma-Delta Modulation (SDM), and Pulse-Width Modulation (PWM) are shown in figure 3.2, where they are used both on images and on regular 1-dimensional signals. For the images, the modulations are slightly different and have other names, though. Different modulations like these three have different properties, which make them suitable for different purposes. For example, SSR is very simple while SDM is more complex but often gives better results. This allows us to trade off between simplicity and signal quality.

Due to the coarse quantization, we might need to use some kind of integration as a way to achieve the desired SNR for the signal. This can be done by oversampling or by parallelization of the circuits. When converting away from the 1-bit domain, low-pass filtering or summation will integrate the values of these extra samples in order to produce high-SNR samples. Note that e.g. PWM can be said to be infinitely oversampled, which explains why it can convey analog signals so well. An interesting consequence of the integration paradigm, is that e.g. a few bit errors in an oversampled bitstream or a few faulty elements in an array of parallel circuits will not cause any big problems. Any such errors in classical digital logic, especially in one of the MSBs, could cause big distortions to the signal. The integration therefore makes the systems more robust.

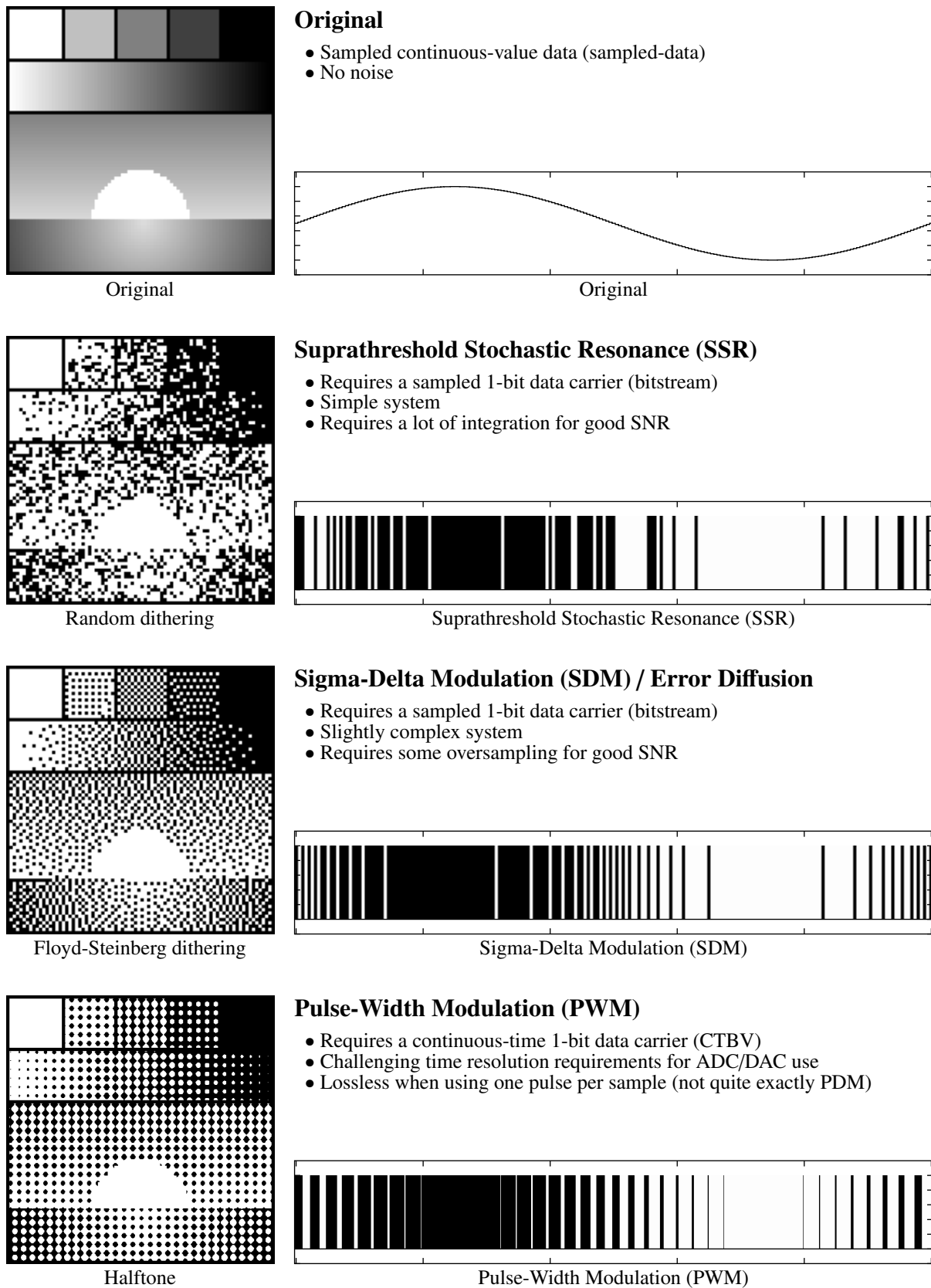


Figure 3.2: Three different 1-bit signal modulations, all of them types of Pulse-Density Modulation (PDM).

Note that no integration is used in figure 3.2, i.e. there are as many 1-bit samples as there are samples in the original signal. However, the original signal has mostly low-frequency components, so when the human vision calculates the average light intensity for different parts of an image or a 1-dimensional plot, this acts as a low-pass filter or, in other words, PDM demodulation. With this way of looking at the setup, the original signals are mostly band-limited, the 1-bit signals are oversampled, and the original signal is recovered when the human vision integrates the 1-bit signals. If we do not take into account that the original signals are band-limited, though, and just look at each sample in isolation, one at a time, the quality of the 1-bit signals will seem much worse than when we do some kind of integration.

An interesting observation to make, is that when the SNR per sample, i.e. the full-band SNR, of an unquantized and possibly oversampled signal is less than about 0 dB, the information content per sample will be less than 1 bit. This helps explain why 1-bit quantization of such a signal, i.e. an SSR system, does not cause much loss. If the information content were many bits, 1-bit quantization would necessarily cause a lot of loss. An SSR system does in fact usually require that there is some noise in the input, i.e. that the SNR per sample is low. An SDM system, on the other hand, does not require this but is still able to perform well. In order to e.g. convey the value 0, an SDM system will rapidly switch between 1 and -1 so that the average is 0. This obviously gives a lot of noise, but by doing the switching in a smart way, the noise can be shifted to higher frequencies, thus mostly avoiding the signal band. So also here the SNR per 1-bit sample is low so that the information content per sample is less than 1 bit, but unlike an SSR system, an SDM system actually adds the noise itself. Either way, there will always be a lot of noise in a 1-bit PDM signal, so in order to get good system performance, our task will be to keep as much of the noise as we can out of the signal band. This gives a high in-band SNR while keeping the full-band SNR low. One way to do this is to simply oversample an SSR system while keeping the noise power constant, so that the noise is spread over a wider band than the signal. Alternatively, the noise Power Spectral Density (PSD) can be kept constant so that oversampling can be used to gather more noise in order to achieve a sufficiently low full-band SNR. This means that it is possible to condition a signal so that 1-bit quantization does not cause much harm to it. A second way to keep the noise out the signal band, is to both oversample and to shape the noise spectrum with an SDM system, as mentioned above. This gives an even better result than an SSR system, but is more complex. Note that parallelized systems will have the same kind of behavior as the oversampled systems discussed here and that continuous-time 1-bit systems can be seen as infinitely oversampled and therefore will have this kind of behavior too. The infinite oversampling does not mean that the noise will be spread over an infinitely wide band, though. The interesting observation to make from all this is that it is the noise, specifically the low full-band SNR, which makes the coarse 1-bit quantization acceptable.

To convert back from a PDM bitstream to analog again, we simply low-pass filter the oversampled 1-bit signal. This gives us an analog signal without the extra out-of-band noise which the oversampled 1-bit signal has. The SNR of the analog signal will be equal to the in-band SNR of the PDM bitstream.

3.3.2 Signal Processing Operations

In addition to using 1-bit signals to represent analog signals as an intermediate format between A/D and D/A conversion, we can also perform signal processing operations on the analog signal while it is still in its 1-bit form. Both sampled and continuous-time 1-bit signals can be processed, and the former case is sometimes called bitstream processing.

There are several potential benefits of using 1-bit signal processing operations. Firstly, we will not need to decimate the signal first if we already have a bitstream from e.g. a sigma-delta ADC. Secondly, the circuits used to perform the processing can sometimes be simpler than corresponding circuits for multi-bit or analog signals. Finally, since 1-bit signals are usually oversampled, we reap some general benefits from oversampling, such as avoiding sharp anti-aliasing and interpolation filters and their large delays, avoiding aliasing of harmonics created by the processing, and getting filters with frequency responses that more closely match analog filters.

The nonlinearity of 1-bit quantization can sometimes be interesting in itself since such an operation can not be performed with linear systems. One use for this is to reject impulsive noise in the input, and another use is in Code Division Multiple Access (CDMA) receivers. When correlating the input signal with a symbol template in such a receiver, a strong interferer would give a large correlation value in an analog system even if the match with the symbol were poor. This is because the correlation value is proportional to the amplitude of the input in this case. In a 1-bit receiver, though, each chip in the symbol is 1-bit quantized before correlation with the template, which means that only the degree of symbol match has anything to say and that the input amplitude will not affect the correlation value.

We can also do signal processing of a more analog character, such as multiplying two analog signals represented by bitstreams by simply XNOR-ing the two bitstreams. This will produce a new bitstream that represents an analog signal which is the product of the two original analog signals.

3.3.3 Overview of Basic Modulation Types, Converters, and Operations

Figure 3.3 gives an overview of a few important underlying data carriers, basic modulation types, modulators / signal converters, and signal processing operations for 1-bit representations of analog signals.

The underlying data carriers and basic modulation types can be seen in the bottom of the figure. Several new ones are shown here which are not shown in figure 3.1, but here we only show those that represent analog signals, not e.g. the one with serialized data bits. By multi-CTBV we simply mean a collection of CTBV signals, and thermometer code means that e.g. the first four of ten CTBV lines are high in order to convey the value four. By classical digital we mean clocked digital logic with multi-bit words, and Pulse-Code Modulation (PCM), which uses this, represents an analog signal by means of a sequence of integers stored in such multi-bit words. Note that the PDM bitstream can be oversampled, but that integration by parallelization is not shown here.

The modulators and signal converters can be seen in the middle of the figure. Their job is to convert the signal back and forth between different basic modulation types without changing

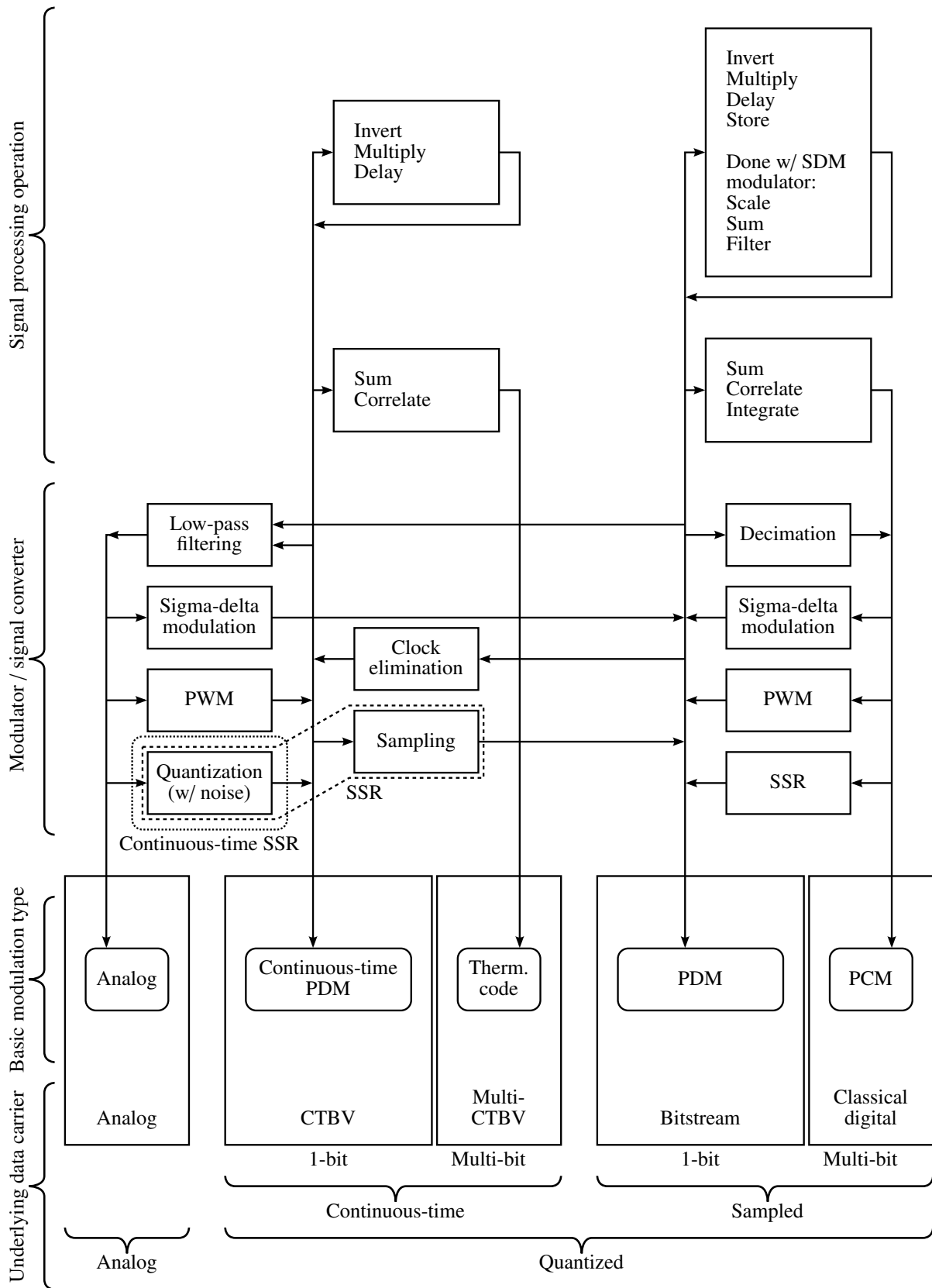


Figure 3.3: Various data carriers, basic modulation types, converters, and operations for 1-bit representations of analog signals.

the analog signal being represented. Conversion from analog to PDM with SSR, SDM, and PWM has already been shown in figure 3.1. In the figure here, though, we show more clearly that continuous-time SSR simply is 1-bit quantization and that normal sampled SSR just adds sampling to this. In general, we can convert any continuous-time PDM signal to a PDM bitstream, i.e. convert a 1-bit signal from continuous-time to sampled form, by simply sampling it. Note that this could cause some aliasing, though. Conversely, we can convert a PDM bitstream to a continuous-time PDM signal by simply eliminating the clock, e.g. by going from transporting the signal with a chain of D Flip-Flops (DFFs) to transporting it with a chain of inverters. The signal will of course still have a discrete-time character, at least until some form of processing changes this. To convert a PDM signal, of either the continuous-time or the sampled type, to an analog signal, we simply low-pass filter it. In addition to converting to and from analog signals like this, it is also useful to be able to convert to and from PCM signals, i.e. normal sampled multi-bit signals. The converters that do this are quite similar to those for analog signals, except that they are implemented digitally. An example of such a converter is a decimator, which converts PDM bitstreams to PCM signals by low-pass filtering and downsampling them. Note that the conversions from PCM to PDM bitstreams usually include interpolation, i.e. an increase in sampling rate. It is easy to see from the figure how SSR and SDM can be used for both multi-bit ADCs and multi-bit DACs. For A/D conversion, we use an SSR or SDM modulator to convert an analog signal to a PDM bitstream, and then we use decimation to convert this to a PCM signal. Conversely, for D/A conversion, we use an SSR or SDM modulator, or even a PWM modulator, to convert a PCM signal to a PDM bitstream, and then we use low-pass filtering to convert this to an analog signal. These approaches to A/D and D/A conversion can be very beneficial due to the simplicity and linearity of the 1-bit ADC and DAC front-ends, as described in section 3.3.1.

The signal processing operations can be seen in the top of the figure. As is shown there, some of them keep the signal within the same basic modulation type, whereas others convert the signal to a different basic modulation type as a result of the processing. Several examples of operations are shown. To invert a signal, we simply use an inverter, and to multiply two signals, we can actually just use an XNOR gate. To delay a signal, we can use a chain of inverters for the continuous-time case and a chain of DFFs for the sampled case. DFFs can also be used to store and replay a signal, but only for the sampled case. By using a sigma-delta modulator in the implementation of the signal processing operations, we are able to perform e.g. scaling, summing, and filtering on PDM bitstreams without changing the signals from 1-bit at the input to multi-bit at the output. Note that there might be other ways to do this which do not require a sigma-delta modulator, though. To perform summing or correlation on PDM bitstreams like this, but with multi-bit output, we simply need to add together some 1-bit values with classical digital logic. If the input is continuous-time PDM signals, it is still possible to perform such summing or correlation, and the output could then be in the form of a continuous-time thermometer code. To integrate a PDM bitstream, i.e. to take the cumulative sum over a long time, or in other words to find the DC value of the analog signal being represented, we can simply use a counter based on classical digital logic. This gives a multi-bit output value.

These examples show some of the rich opportunities for conversion and processing of 1-bit representations of analog signals. With possible benefits such as better linearity, simpler circuits, native processing of input from 1-bit ADCs, some immunity to bit errors, and more, this makes

1-bit signal processing a very interesting option when designing systems.

3.4 Power Consumption

One of the benefits of using 1-bit signal processing in system designs is that in some cases, these systems might have a lower power consumption than corresponding analog or multi-bit systems. This is true not just for SDM systems but also for SSR systems, despite their simplicity. Low power consumption is always a desirable property and could help tip the balance in favor of 1-bit signal processing. The power reduction is explained by several different mechanisms, and depending on the specific signal processing task at hand, one or more of these mechanisms might be in effect. We will here discuss the different mechanisms.

This section is organized as follows: In section 3.4.1 and 3.4.2, we investigate how much energy is required to convey a sample with a given dynamic range through a system. We look at native samples here, i.e. we do not use any kind of modulation/encoding. In section 3.4.3, we compare the power consumption of analog, multi-bit, and 1-bit modulations/encodings. This analysis is based on the results about native samples. The first two sections give a level of detail that is not very important for the comparison of system types in the third section, so if only the comparison is of interest, these two sections can be skipped. They still give some valuable insights into fundamental power consumption trade-offs, though. In section 3.4.4 and 3.4.5, we look into a few other types of power consumption benefits we get from switching from analog to digital systems and from switching from multi-bit to 1-bit systems.

3.4.1 Unit-Element Amplifier

In order to make a general analysis of the power consumption and dynamic range of electronic systems, we will introduce a simple unit-element amplifier and use this as the basis of our analysis. All the amplifier does is to convey the signal at unity gain. It does not perform any form of signal processing operations. This simple setup allows us to concentrate on the power consumption of the amplifier and the thermal noise it adds to the signal. In this section and section 3.4.2, we will concern ourselves with native signal values without any modulations or encodings, i.e. simple analog values. This includes digital bits, though, since also these are analog values electrically speaking. Multi-bit encoding of signals will be discussed in section 3.4.3.

The unit element is a unit both in space and time, i.e. we can have several parallel amplifiers and they are only powered up for the duration of each time unit. By doing signal integration over several units, we can easily increase the dynamic range of the system at the cost of increased power consumption. This is why we base our analysis on a unit element. The integration of the units namely gives us a degree of freedom for trade-off between dynamic range and power consumption which gives insight into the relationship between the two. In addition to this, we will also get a clear overview of the different ways to perform such integration. One way to perform the integration is to use parallelization. To do this, we simply use larger gate widths in the transistors. We do not need to actually have many separate amplifiers which we integrate the results from, i.e. we do not need to use the unit-element design methodology in the actual circuit implementation. A different way to perform the integration is to use oversampling. To do this, we simply integrate the result from several units over time, i.e. several samples. This can be done e.g. with a low-pass filter. The effect of oversampling is essentially the same as for parallelization, but there are a few differences, though, which we discuss further in section 4.6. In order to compare analog and sampled systems, we will sometimes consider analog systems as sampled systems with high oversampling, or in other words, with adjoining units in the time dimension. This provides us with an important realization about analog systems, namely that the continuous operation is in fact a form of integration.

The circuit model we will use for the unit element is a simple analog amplifier with unity gain and output impedance R_u . The output of the amplifier connects to the input of another amplifier, and we assume that the input impedance of the amplifier is purely capacitive with a capacitance of C_u . The output of the amplifier is therefore essentially connected to a low-pass RC filter. The details of the amplifier will be discussed later. Parallelized versions of the unit-element amplifier have $R = R_u/n$ and $C = C_u n$, where n is the integration factor. We will in the following mostly use R and C instead of R_u and C_u since they are more succinct terms, i.e. we will analyze the parallelized amplifier rather than the unit-element amplifier. These two amplifiers are equivalent if we choose $n = 1$.

Our analysis will give some insights into the fundamental limits of power consumption. Real systems will of course be a bit more complicated than the simplified model we use here, though. For one thing, we will only analyze the overall electrical behavior of the unit elements. We will not look into transistor models or specific circuit topologies. Despite this, we should still be able to learn something about how the different quantities involved depend on each other, and we should also be able to find reasonable estimated values for e.g. the power consumption. Note

that a long cascade of amplifiers would cause a lot of noise to be added to the signal since each amplifier would contribute noise. The dynamic range of such a cascade will therefore be lower than for the single amplifier which our analysis will focus on. This must be kept in mind for systems with long signal processing chains.

In the current section, we will find the power consumption of analog systems, i.e. the power consumption of the amplifier during continuous operation. After that, in section 3.4.2, we will look into sampled systems, i.e. systems where the amplifier is powered up only for a short amount of time for each sample. The latter section also includes other extensions of the analysis.

3.4.1.1 Power Consumption During Continuous Operation

The power consumption of the unit-element amplifier used in an analog system, i.e. with the amplifier always powered up, can depend on e.g. the amplifier type, the signal distribution, the signal spectrum, and the output load. In order to simplify the analysis, we will introduce an ideal amplifier without these dependencies and simply state that the real unit-element amplifier has a power consumption P_u given by the product of the power consumption $P_{u\text{-ideal}}$ of this reference amplifier and a relative power consumption factor P_r . We can then skip the complex analysis needed to find out what the exact value of P_r is, since this value is not that important when we just want to look at the big picture. We do the same for the parallelized amplifier. This gives us $P_u = P_{u\text{-ideal}} \cdot P_r$ and $P = P_{\text{ideal}} \cdot P_r$.

According to the maximum power theorem [W], the amplifier will have maximum power output when the load is equal to the output impedance, i.e. when the load is R . When we also have that the Root Mean Square (RMS) amplitude of the signal is $v_{\text{signal-max}}$, the power delivered to the load will be $(v_{\text{signal-max}}/2)^2/R = v_{\text{signal-max}}^2/4R$. The ideal amplifier will consume exactly this amount of power since it will have an efficiency of 100 %. We therefore define $P_{\text{ideal}} = v_{\text{signal-max}}^2/4R$. The maximum power theorem ensures that the power output will never exceed this power consumption value. When the power output is lower than this due to a different load, though, we define the power consumption of the ideal amplifier to be the same as before since this simplifies the analysis. In other words, the unused part of the output power will be burned off inside the amplifier. This gives us $P_u = P_{u\text{-ideal}} \cdot P_r = (v_{\text{signal-max}}^2/4R_u) \cdot P_r$ for the unit-element amplifier and $P = P_{\text{ideal}} \cdot P_r = (v_{\text{signal-max}}^2/4R) \cdot P_r$ for the parallelized amplifier. These and other equations are listed in figure 3.4.

The ideal amplifier behaves a lot like a transmission line with characteristic impedance R and signal voltage $v_{\text{signal-max}}/2$. With a load R , the output voltage will be $v_{\text{signal-max}}/2$, and without any load, the output voltage will be $v_{\text{signal-max}}$ since the voltage of the reflected signal will be added to the voltage of the incident signal. The power transmitted through the transmission line will be $(v_{\text{signal-max}}/2)^2/R = v_{\text{signal-max}}^2/4R$, and the unused output power will be reflected back through the transmission line. Because they behave the same and because the behavior of the transmission line is well-known, it might be easier to think of the ideal amplifier as a transmission line instead of as an amplifier with all the specific properties described above.

Note that the signal is bipolar, i.e. that it swings both above and below ground. If the peak amplitude is e.g. three times the RMS amplitude, this means that the amplifier must have a signal value range of $6v_{\text{signal-max}}$. Also note that we define $v_{\text{signal-max}}$ as the maximum signal level which does not cause distortion, thus making it suitable for use in expressions for dynamic range. This

means that the actual signal level will not necessarily be equal to $v_{\text{signal-max}}$. Signals that are weaker than this will therefore have an SNR which is lower than the dynamic range of the system. Such signals might also cause the amplifier to consume less power. We will here generally assume that the signal is at its maximum level, though.

3.4.1.2 Real Amplifiers Compared to the Ideal Amplifier

We use the ideal amplifier as a stand-in for real amplifiers, but exactly how close is the actual power consumption of a real amplifier to the power consumption of the ideal amplifier? In other words, how close is P_r to 1?

With a load R , the ideal amplifier will have a power consumption equal to the power delivered to the load. A real amplifier will have to deliver the same amount of power to the load, so it will consume at least as much power as the ideal amplifier, i.e. $P_r \geq 1$. The power consumption of the real amplifier will probably not be too many times the power consumption of the ideal amplifier, though.

With a load C , we can in some cases get $P_r \ll 1$, for example in a digital system where the signal rarely changes. The amplifier is e.g. an inverter in such a system, so as long as the signal does not change, the amplifier will not consume any power. This gives a low average power consumption. However, if the signal is a random bitstream with a 50 % chance for either bit value and no correlation between the samples, we will get $P_r \approx 2$, which will be shown below.

The observations we have made here demonstrate that the ideal amplifier is a reasonable choice of reference for the real amplifiers, also for capacitive loads even though the ideal amplifier was created with resistive loads in mind. Specifically, we have seen that the power consumption of real amplifiers is close to the power consumption of the ideal amplifier in most cases. For the overall analysis we are trying to perform here, we are mostly interested in how the power consumption scales with the dynamic range, the sampling frequency, etc., so any differences in power consumption which are more or less constant are not that important for this purpose.

3.4.1.3 Three Distinct Energy Quantities

Before we investigate the amplifier any further, we need to make a clear distinction between three different energy quantities which can easily be confounded otherwise.

The first quantity is average energy consumption per sample. This is the energy used by the amplifier from the power supply.

The second quantity is average energy output per sample. This is the energy which the amplifier delivers to the load. It is here important to distinguish between results based on measurements of the real power, with unit W, and results based on other metrics such as apparent power, with unit VA. With reactive loads, the results can be very different for different metrics.

The third quantity is average energy per degree of freedom in the output signal. By output signal we here mean the signal at the output node of the amplifier, i.e. after the output impedance, or, in other words, the signal on the load. Any degree of freedom in that output signal, such as a sine or cosine amplitude or the voltage of a sample, will have an average thermal noise energy of $kT/2$ and a signal energy which is proportional to that value. This property can be observed not only with resistive loads but also with capacitive loads. Since the average energy

output to a capacitor is zero when we measure the real power and frequency-dependent when we measure the apparent power, it is clear that the degree-of-freedom energy quantity is something quite different than the energy-output quantity. This interesting quantity will be discussed further below.

For the ideal amplifier and a load R , all three energy quantities will be the same. First of all, the energy consumption will be the same as the energy output due to the 100 % efficiency of the ideal amplifier and the matched load. Secondly, the voltage over the load resistor for a given sample is a degree of freedom and will therefore have an average energy proportional to $kT/2$. The amplifier needs to output exactly this amount of energy during the sample by basically holding the voltage at the given level for the duration of the sample. The degree-of-freedom energy quantity is therefore equal to the energy-output quantity in this case.

For a real amplifier and a load C , which is the setup that we are actually interested in, the energy quantities will no longer be equal to each other. First of all, when switching to a real amplifier but still using a load R , only the energy consumption will change since the two other energy quantities are unaffected by amplifier architecture. In other words, P_r will no longer necessarily be equal to 1. Secondly, when switching to a load C for the real amplifier, the value of C , the signal frequency, etc. could start affecting the energy consumption of the amplifier. In other words, it will become more difficult to predict the value of the energy consumption quantity. Another consequence of switching the load to a capacitor is, as mentioned above, that the average energy output will be zero when we measure the real power and frequency-dependent when we measure the apparent power. This means that also the output-energy quantity will change and become more difficult to predict. The energy per degree of freedom will not change, though, as it always stays proportional to $kT/2$.

3.4.1.4 The Ideal Amplifier Gives Exact Energy per Degree of Freedom

As stated earlier, the ideal amplifier with a load R has an energy consumption per sample which is equal to the energy per degree of freedom. Furthermore, the energy per degree of freedom is independent of amplifier architecture and of whether the load is R or C . Lastly, the power consumption of the ideal amplifier is independent of the load. All of this means that the average energy consumption per sample of the ideal amplifier for the maximum signal level, E_{ideal} , will always be equal to the average energy per degree of freedom in the output signal for the maximum signal level, $E_{\text{dof-signal-max}}$, even in the output from e.g. a real amplifier with a load C , i.e. $E_{\text{ideal}} = E_{\text{dof-signal-max}}$.

Another way to put this is that the ideal amplifier does not only give us an estimate of the power consumption of real amplifiers but also an exact value for the energy per degree of freedom. We can therefore see P_r as specifying how many times bigger the energy consumption per sample of a real amplifier is compared to the energy per degree of freedom such as the output voltage for a sample.

3.4.1.5 Proof of $kT/2$

As stated above, the average energy per degree of freedom in the output signal, i.e. the signal on the load, is proportional to $kT/2$. This holds for degrees of freedom such as sine and cosine

amplitudes, i.e. the real and imaginary parts of frequency component phasors, and the voltage of samples. The property can be observed both with resistive and capacitive loads, but note that the energy is measured in slightly different ways in the two cases, though. We will now provide proof for all of these claims.

The average thermal noise energy per degree of freedom is exactly $kT/2$, i.e. there is no scaling factor for the thermal noise. The average signal energy per degree of freedom for the maximum signal level, $E_{\text{dof-signal-max}}$, however, scales with the dynamic range of the signal as measured in the given degree of freedom, DR_{dof} . The relationship is simply $E_{\text{dof-signal-max}} = \text{DR}_{\text{dof}} \cdot kT/2$. This is what we mean with the expression “proportional to $kT/2$ ”. Note that by using the equalities from earlier, we can now easily show that the energy consumption per sample will be $E_{\text{ideal}} = E_{\text{dof-signal-max}} = \text{DR}_{\text{dof}} \cdot kT/2$ for the ideal amplifier and $E = \text{DR}_{\text{dof}} \cdot (kT/2) \cdot P_r$ for real amplifiers.

With a load R , i.e. a resistive load matched to the output impedance of the amplifier, we measure the average energy per degree of freedom in the output signal as the energy dissipated on the load resistor due to the value of the given degree of freedom. Note that a transmission line with impedance R will behave like the load resistor R in many ways here, so the results we get for a load R can also be used to analyze the energy per degree of freedom of a signal in a transmission line.

The PSD of the thermal noise delivered from one resistor to another resistor with the same value, is kT . This power can be said to be divided between a sine and a cosine component for each frequency, so the PSD will thus be $kT/2$ for each of these two components. If we take a DFT of e.g. one second of signal, we will get one frequency component for each hertz of the spectrum. To get the energy of one of these frequency components, we can therefore multiply the PSD with 1 Hz to get the power of the sinusoid and then multiply with 1 s to get the total energy. We then get $(kT/2) \cdot 1 \text{ Hz} \cdot 1 \text{ s} = kT/2$, a result which we would also get for other signal lengths than 1 s. In other words, the amplitude of a sine or cosine component, which is a degree of freedom, has energy $kT/2$ that gets dissipated on the load resistor. This quantity matches our definition of average energy per degree of freedom, so we have now shown that each degree of freedom in the spectrum has an average thermal noise energy of $kT/2$.

In addition to this consideration of the frequency domain, it is also possible to look at the voltage level of each sample in the time domain as a degree of freedom. Since the load is matched to the output impedance of the amplifier, the thermal noise delivered from the amplifier to the load will have a PSD of kT . This means that if we sample the output signal with sampling rate f_s and a brick-wall anti-aliasing filter, the noise power in the sampled signal will be $kT \cdot (f_s/2)$ and the noise energy per sample will be $kT \cdot (f_s/2) \cdot T_s = kT/2$. This shows that also samples are degrees of freedom with an average thermal noise energy of $kT/2$ each.

With a load C , we measure the average energy per degree of freedom in the output signal as the average energy stored in the capacitor due to the value of the given degree of freedom. Note that this energy is not dissipated by the load, it is just stored there temporarily. Each sample of the output signal will have an RMS thermal noise voltage $v_n = \sqrt{kT/C}$, as stated earlier. The average thermal noise energy on the capacitor will therefore be $\frac{1}{2}Cv_n^2 = \frac{1}{2}C(\sqrt{kT/C})^2 = kT/2$. This shows that also with a capacitive load, the samples are degrees of freedom with an average thermal noise energy of $kT/2$ each. The frequency components seem to be a bit more difficult to

work with for capacitive loads, though, so we will not look into those here.

3.4.1.6 Remaining Questions About $kT/2$

We have shown for a few examples that the average thermal noise energy per degree of freedom over the load is $kT/2$. We have not stated a general rule about this, though, nor will we try to do so here. Part of the reason for this is that some questions still remain about the general properties of the rule.

One question is how general the rule is, i.e. for which circuit configurations it applies. As an example, a resistive load which is not matched to the source resistance will not receive thermal noise power with a PSD of kT , so in this case the rule does not seem to apply, at least in the form we have used so far.

A second question is how the energy should be calculated in general. As we have seen, the energy per degree of freedom is calculated differently for resistive and capacitive loads. This suggests that there is a general rule for how the energy per degree of freedom should be calculated, and that this general rule results in different methods for different types of loads. It would also be interesting to get some more insight into exactly what this kind of energy really is, i.e. what does this energy quantity actually mean.

A third question is why there exists such a $kT/2$ rule at all. So far we have reached the $kT/2$ result by simply making derivations based on the thermal noise PSD of kT , but could there be a deeper reason behind the rule than this?

The equipartition theorem [W] seems to be the answer to the last question. It states something like that particles in a system in thermal equilibrium have an average energy of $kT/2$ per degree of freedom. These degrees of freedom could be e.g. velocity components, which give kinetic energy, or position components, which could give potential energy. In the context of thermal noise on capacitors, it has been stated like this: “Any system in thermal equilibrium has state variables with a mean energy of $kT/2$ per degree of freedom.” [Wiki 15 ♻]. So the theorem seems to hold not only for tiny particles in a system, but also for macroscopic degrees of freedom such as the voltage on a capacitor. The voltage over a resistor is perhaps less of a state variable than the voltage on a capacitor, but as we have shown, the $kT/2$ result seems to hold also for resistors. Exactly why that is, could be interesting to look into, though.

The answers to all of these questions can perhaps be found in literature, but we will not look any further into this here.

3.4.2 Unit-Element Amplifier in Sampled Systems

In the previous section, we explored the power consumption of analog systems, i.e. the power consumption of the unit-element amplifier during continuous operation. In the current section, we will use what we found there to investigate the power consumption of sampled systems, i.e. systems where the amplifier is powered up only for a short amount of time for each sample. A common type of such sampled systems are those that are based on digital circuits. We will here provide a detailed analysis of the digital inverter since this is an often-used basic building block in these systems and since other digital circuits share many properties with it. What we find out about the inverter will therefore tell us a lot about the behavior of complete digital systems, and the analysis will thus allow us to e.g. estimate the power consumption of such systems. In addition to investigating these sampled systems, we will also look a bit more into the analog systems we started with.

Power consumption is the central theme of our analysis, but in this section we will also look into related concepts such as signal integration and the dynamic range of the systems.

As mentioned earlier, we will only consider native samples here, i.e. samples without any form of modulation or encoding. This is important here since we will be working with the dynamic range of the samples and since modulation or encoding would therefore have complicated the analysis. One example of where native samples are used is in sampled-data systems, which have analog sample values, and another example is for digital bits. In the latter case, we only consider the underlying analog values of the bits, not the logical values themselves. An example of where non-native samples are used is in systems with multi-bit sample values. The reason why these samples are not native is that they utilize value encoding and are therefore not native according to the criteria we just gave.

Figure 3.4 shows a list of equations relating to the dynamic range and power consumption of the unit-element amplifiers, both with and without the use of integration. In the following sections, we will explain these equations.

3.4.2.1 Integration by Parallelization

As already mentioned, the unit-element amplifier has output impedance R_u and input load C_u . By integrating through parallelization, i.e. by using larger gate widths, and by using an integration factor n , we get a new and bigger amplifier with values $R = R_u/n$ and $C = C_u n$.

3.4.2.2 Frequencies and Related Quantities

The low-pass RC filter we get between two cascaded unit-element amplifiers has a cut-off frequency $f_c = 1/2\pi R_u C_u = 1/2\pi RC$, which we can see is independent of n . The noise from the amplifier gets filtered by this filter, which gives an equivalent noise bandwidth $BW_{\text{noise}} = f_c \cdot \pi/2 = 1/4RC$. By assuming that the signal and noise are filtered by a brick-wall filter with this bandwidth, we can easily find the native sampling rate $f_{\text{su}} = 2 \cdot BW_{\text{noise}} = 1/2RC$ of the unit element. With this assumption and this sampling rate, the unit element samples will have independent noise and the noise will be band-limited, which makes the analysis easier to perform. The sampling period $T_{\text{su}} = 1/f_{\text{su}} = 2RC$ follows from f_{su} .

The sampling rate that the system actually uses is named f_s , and we assume that these samples will have independent noise, too. f_s must not be larger than f_{su} . An analog system is modeled with $f_s = f_{su}$.

The signal of interest, i.e. the Nyquist-rate signal, has a sampling rate of $f_{s\text{-Nyquist}}$. If the signal is analog, we define $f_{s\text{-Nyquist}}$ as the sampling rate that would be required to sample it, i.e. the Nyquist rate of the signal.

$\text{OSR}_u = f_{su}/f_{s\text{-Nyquist}} = 1/2RCf_{s\text{-Nyquist}}$ gives the native oversampling ratio of the unit elements with respect to the sampling rate of the signal of interest, and $\text{OSR} = f_s/f_{s\text{-Nyquist}}$ gives the actual oversampling ratio used in a sampled system.

3.4.2.3 Signal, Noise, and Dynamic Range

The signal strength is given by $v_{\text{signal-max}}$, which is the maximum RMS amplitude that the signal can have without causing too much distortion. This quantity is discussed further in section 3.4.1.1.

The thermal noise in a unit element sample, or on the analog output node in general, is so-called kTC noise [W] because of the configuration with R_u and C_u . The RMS amplitude of this noise is $v_{nu} = \sqrt{4kTR_u \cdot BW_{\text{noise}}} = \sqrt{kT/C_u}$, where k is the Boltzmann constant and T is the temperature in kelvin. An amplifier with integration by parallelization has less thermal noise, $v_n = v_{nu}/\sqrt{n} = \sqrt{4kTR \cdot BW_{\text{noise}}} = \sqrt{kT/C}$, because of the increased capacitance. When integrating by oversampling, the noise voltage stays the same, but the noise is spread over a wider band. The RMS amplitude of the in-band noise or in other words the noise in the Nyquist-rate samples is therefore less than v_n , i.e. less than the noise without oversampling. Specifically, the amplitude is $v_{n\text{-Nyquist}} = v_{nu}/\sqrt{n \cdot \text{OSR}} = \sqrt{4kTR \cdot BW_{\text{noise}}}/\sqrt{\text{OSR}} = \sqrt{kT/C}/\sqrt{\text{OSR}}$.

From these equations, it follows that the dynamic range of a unit element sample is $\text{DR}_u = v_{\text{signal-max}}^2/v_{nu}^2 = \frac{1}{2}C_u v_{\text{signal-max}}^2/(kT/2)$, that a parallelized amplifier sample has $\text{DR} = \text{DR}_u \cdot n = v_{\text{signal-max}}^2/v_n^2 = \frac{1}{2}C v_{\text{signal-max}}^2/(kT/2)$, and that a Nyquist-rate sample in an oversampled system has $\text{DR}_{\text{Nyquist}} = \text{DR}_u \cdot n \cdot \text{OSR} = v_{\text{signal-max}}^2/v_{n\text{-Nyquist}}^2 = \frac{1}{2}C v_{\text{signal-max}}^2 \cdot \text{OSR}/(kT/2)$.

3.4.2.4 Power Consumption During Continuous Operation

The power consumption of the unit-element amplifier during continuous operation was derived in section 3.4.1.1. For convenience, we mention the results again here, namely that $P_u = P_{u\text{-ideal}} \cdot P_r = (v_{\text{signal-max}}^2/4R_u) \cdot P_r$ and that $P = P_{\text{ideal}} \cdot P_r = (v_{\text{signal-max}}^2/4R) \cdot P_r$.

3.4.2.5 Energy Consumption per Sample

Since we know what the power consumption of the amplifiers is during continuous operation, namely P_u and P , we can easily find the energy consumption per sample when these amplifiers are used in sampled systems. We assume these systems are able to power up the amplifiers for short amounts of time, the unit element sampling interval T_{su} , at a sampling rate of f_s . We further assume that this is enough time for a sampled-data system to settle, and that this is the time a digital circuit uses to switch levels. The energy consumption per sample for a unit-element amplifier, E_u , and for a parallelized amplifier, E , can then be calculated by multiplying the always-on power consumption with T_{su} . This gives us $E = P \cdot T_{su} = \frac{1}{2}C v_{\text{signal-max}}^2 \cdot P_r$, which

we can also write as $E = DR \cdot (kT/2) \cdot P_r$. An oversampled system uses OSR samples per Nyquist-rate sample of the signal, so the energy consumption per Nyquist-rate sample in such a system will be $E_{\text{Nyquist}} = E \cdot \text{OSR}$.

As expected, the RMS amplitude of the maximum signal level, $v_{\text{signal-max}}$, the power consumption overhead factor for the real amplifier, P_r , and the oversampling ratio OSR are all involved in the expressions for energy consumption per sample. Other than this, it is only C which determines the result. The energy in a capacitor is given by $\frac{1}{2}CV^2$, so as we can see, the energy consumption per sample is mainly determined by the energy that would be required to charge up the capacitive load to $v_{\text{signal-max}}$. Note that it is not as simple as that the energy consumed by the ideal amplifier for a sample will be directly transferred to an empty capacitor. The reason for this is that the capacitor will already have a voltage level due to the previous sample and that for arbitrary signals, the ideal amplifier will not have a 100 % energy transfer to loads other than R .

The role of C in the expression for energy consumption means that the smaller the capacitive load is, the lower the energy consumption per sample will be. However, as can be seen in the expressions for dynamic range, a smaller capacitance would not only reduce the energy consumption but also the dynamic range. The reduction of C is therefore in fact a trade-off rather than just a matter of circuit optimization.

The output impedance R of the amplifier, on the other hand, does not have anything to say for the energy consumption per sample. In other words, a powerful amplifier with a lower R and the same C would not be more energy consuming than a weak amplifier. The reason for this is that a powerful amplifier would charge up the capacitive load quicker, so that it would need to be powered up for a shorter amount of time per sample than a weak amplifier. The maximum sampling rate f_{su} would be higher for such an amplifier, though, so reducing R does have some advantages.

3.4.2.6 Energy Consumption of Inverters

Although it is difficult to find the exact value of P_r in general, it is easy to find a good estimate for it in the case of digital systems. In such systems, the signal switches between 0 V and V_{DD} . If we refer the signal to $V_{\text{DD}}/2$, we get $v_{\text{signal-max}} = V_{\text{DD}}/2$ and thus signal levels $V_{\text{DD}}/2 \pm v_{\text{signal-max}} = V_{\text{DD}}/2 \pm V_{\text{DD}}/2$, i.e. 0 V and V_{DD} . When the signal switches levels, the charge $Q = CV_{\text{DD}}$ flows through the capacitive load. If we assume the capacitor is connected to ground on the other side, a 0-to-1 transition performed by the PMOS of an inverter will consume the energy $E = QV_{\text{DD}} = CV_{\text{DD}}^2$. Half of this energy, $\frac{1}{2}CV_{\text{DD}}^2$, will be stored in the capacitor and the other half will be dissipated in the transistor. A 1-to-0 transition does not consume any energy since the energy in the capacitor is then simply dissipated in the NMOS of the inverter without drawing any current from the power supply. If the capacitor is connected to V_{DD} instead of ground, the 1-to-0 transition will consume energy instead of the 0-to-1 transition. If the capacitive load is $C/2$ to ground and $C/2$ to V_{DD} , each transition will have half the energy consumption, which means that the total for both transitions will still be the same. Also with a $\pm V_{\text{DD}}/2$ power supply and the capacitor connected to ground, the total energy consumption will be the same. The energy consumption is therefore always CV_{DD}^2 for the two transitions combined.

If the signal is a random bitstream with a 50 % chance for either bit value and no correlation between the samples, each of these transitions happen in average every fourth sample. When the

bit value is the same as for the previous sample, no energy is consumed. The average energy consumption per sample will therefore be $CV_{DD}^2/4$. Since we have $v_{\text{signal-max}} = V_{DD}/2$, we get $CV_{DD}^2/4 = Cv_{\text{signal-max}}^2 = \frac{1}{2}Cv_{\text{signal-max}}^2 \cdot 2$. The general expression for average energy consumption per sample is $E = \frac{1}{2}Cv_{\text{signal-max}}^2 \cdot P_r$, which means that $P_r = 2$ in this case. In other words, an inverter uses twice as much power as the ideal amplifier when the signal is a random bitstream. In practice, a short-circuit current might flow through the inverter for a short period of time when both transistors are on during switching. This will increase P_r . The switching might also often happen before the signal level has settled to the rail voltage. This will decrease P_r . Because of these effects, we will instead conclude more cautiously that an inverter has $P_r \approx 2$. Keep in mind that this result is for signals that are random as described above.

3.4.2.7 Discussion of the Expression for Energy Consumption

As shown above, the average energy consumption per sample is given by $E = DR \cdot (kT/2) \cdot P_r$. We will now summarize and discuss the meaning and implications of this equation.

$kT/2$ is the average thermal noise energy per degree of freedom, i.e. the average thermal noise energy on the capacitor. The thermal noise voltage over the capacitor is $\sqrt{4kTRB} = \sqrt{kT/C}$ where $B = BW_{\text{noise}} = 1/4RC$ is the equivalent noise bandwidth of the RC low-pass filter. The thermal noise energy on the capacitor is therefore $\frac{1}{2}C(\sqrt{4kTRB})^2 = \frac{1}{2}C(\sqrt{kT/C})^2 = kT/2$.

$DR \cdot (kT/2)$ is the average signal energy per degree of freedom when the RMS amplitude of the signal is $v_{\text{signal-max}}$, i.e. the average signal energy on the capacitor. In other words, the ratio between the signal energy and the thermal noise energy on the capacitor determines the dynamic range of the system. The signal energy on the capacitor is therefore $\frac{1}{2}Cv_{\text{signal-max}}^2 = DR \cdot (kT/2)$.

$DR \cdot (kT/2)$ is also the average energy consumption per sample of the ideal amplifier we have defined for ourselves. However, even though the energy consumption of the amplifier is equal to the capacitor energy, there is no direct energy transfer going on here. This was explained earlier. It is useful to note that the ideal amplifier in many ways is equivalent to a transmission line with thermal noise voltage $\sqrt{4kTRB}/2 = \sqrt{kT/C}/2$, signal voltage $v_{\text{signal-max}}/2$, thermal noise energy $kT/2$ per sample, and signal energy $DR \cdot (kT/2)$ per sample. This equivalence makes it easier to understand the ideal amplifier. The bandwidth-dependence of the thermal noise might cause some confusion, but note that e.g. a doubling of the sampling rate will both double the bandwidth B and thus the thermal noise power, and also cut the sampling interval in half. Since energy is power multiplied by time, this explains why the average thermal noise energy per sample is $kT/2$ regardless of what the sampling rate or B is.

$DR \cdot (kT/2) \cdot P_r$ is the average energy consumption per sample of a real amplifier. P_r specifies how much power a given real amplifier uses compared to the ideal amplifier, but its value can be difficult to estimate and might depend on many factors. Even the properties of the signal could have much to say. P_r is probably not excessively far away from unity in most cases, though.

There are expressions for the energy consumption per sample both for no integration, integration by parallelization, and integration by oversampling, or in other words, expressions based on both DR_u , DR , and DR_{Nyquist} . All of these are on the form $E_X = DR_X \cdot (kT/2) \cdot P_r$, i.e. the average energy consumption per Nyquist-rate sample and the dynamic range of those samples have the same relationship no matter how integration is performed. The important lesson here is therefore that energy consumption is proportional to the dynamic range, i.e. that a large dynamic

range always costs a large amount of energy. If we want to reduce the energy consumption while keeping the dynamic range at a given value, the only parameters we can tune are T and P_r . It is difficult to do much about these, though.

Note that all the results mentioned here are for native samples. By encoding the values as e.g. digital integers, it is possible to achieve arbitrarily large dynamic ranges without consuming much energy.

3.4.2.8 Overall Power Consumption

Now that we know the average energy consumption per sample, we can easily calculate the average overall power consumption P_{avg} for a signal with a given sampling rate $f_{s\text{-Nyquist}}$. We can find this power consumption by e.g. multiplying the energy consumption per sample of the parallelized amplifier, E , with the oversampled sampling rate f_s . We then get $P_{\text{avg}} = E \cdot f_s = \frac{1}{2} C_u v_{\text{signal-max}}^2 \cdot n \cdot \text{OSR} \cdot f_{s\text{-Nyquist}} \cdot P_r$. The first part, $\frac{1}{2} C_u v_{\text{signal-max}}^2$, is the energy consumption per sample of the ideal unit-element amplifier without any integration. This is multiplied with the integration factors n and OSR to give us the total energy consumption per Nyquist-rate sample. Further multiplying this with the Nyquist rate $f_{s\text{-Nyquist}}$ gives us the average power consumption of the ideal amplifier. By multiplying with the final factor P_r , we get the average power consumption of the real amplifier.

The overall power consumption can also be given as $P_{\text{avg}} = \text{DR}_{\text{Nyquist}} \cdot (kT/2) \cdot f_{s\text{-Nyquist}} \cdot P_r$. As discussed earlier, the average energy consumption per Nyquist-rate sample is $\text{DR}_{\text{Nyquist}} \cdot (kT/2) \cdot P_r$. By multiplying this with the Nyquist rate $f_{s\text{-Nyquist}}$, as above, we will also here get the average power consumption of the real amplifier.

These expressions for P_{avg} can be useful when estimating the power consumption of systems, and they also give some insights into the fundamental connections and trade-offs relating to power consumption.

3.4.2.9 The Effects of Tuning the Different Parameters

With all the equations in place, we can now take a look at what happens to the power consumption when we tune the different parameters. The dynamic range will be exactly proportional to the power consumption if we assume that T , P_r , and $f_{s\text{-Nyquist}}$ are constant, so there is no need to consider the effects on the dynamic range separately.

If we e.g. halve R_u , the power consumption of analog systems will be doubled, as can be seen in the expression for P , but the power consumption of sampled systems will be unaffected, as can be seen in the expression for P_{avg} . The reason for this difference is that the maximum sampling rate f_{su} will be doubled and thus the minimum sampling interval T_{su} will be cut in half. In the sampled systems, the amplifier is powered up for the time interval T_{su} for each sample, at a sampling rate of f_s . Since the power consumption of the amplifier is doubled and its duty cycle halved, the average power consumption will stay the same for such sampled systems. The analog systems, on the other hand, can be said to be sampled systems with $f_s = f_{\text{su}}$ and thus $\text{OSR} = \text{OSR}_u$, i.e. with the samples so close together that the amplifier will always be on. When f_{su} is doubled, OSR in the expression for P_{avg} will be doubled too, which explains why P_{avg} will be doubled for analog systems.

If we double C_u , the power consumption of sampled systems will be doubled and the power consumption of analog systems will be unaffected, i.e. the opposite of above. Also here we can see this from the expressions for P and P_{avg} . In this case, f_{su} will be halved and thus T_{su} doubled, though. Similar to above, we can therefore explain the doubling of the power consumption of sampled systems with the doubling of $P \cdot T_{\text{su}}$ and the unaffected power consumption of analog systems with the simultaneous doubling of C_u and halving of OSR in the expression for P_{avg} .

If we double n , thus halving R and doubling C , or in other words if we perform integration by parallelization, we will get the effects from both cases above. This means that the power consumption of both analog and sampled systems will be doubled and that f_{su} will be unaffected.

If we double OSR, i.e. if we perform integration by oversampling, the power consumption of sampled systems will be doubled, as can be seen in the expression for P_{avg} . In analog systems, though, we can not tune OSR freely since $\text{OSR} = \text{OSR}_u$, which is given by R , C , and $f_{\text{s-Nyquist}}$. In sampled systems, where OSR can be tuned, the choice of value for OSR is limited by the inequation $f_{\text{s-Nyquist}} \cdot \text{OSR} \leq f_{\text{su}}$.

If we reduce the supply voltage V_{DD} , several different changes will occur. $v_{\text{signal-max}}$ will be reduced, R_u will increase due to the lower gate voltage, and there might also be a change in the equivalent value of C_u due to the voltage dependence of the gate capacitance. Furthermore, P_r might also be affected if e.g. the $v_{\text{signal-max}}/V_{\text{DD}}$ ratio changes. The result of all this will most likely be that both the power consumption and the maximum sampling rate f_{su} are reduced.

3.4.2.10 Calculated Results for CMOS Inverters

In CMOS technology nodes around 90 nm, minimum-size inverters have an input capacitance on the order of $C_u = 1$ fF. This gives a noise voltage of $v_{\text{nu}} = \sqrt{kT/C_u} = 2$ mV and, with $V_{\text{DD}} = 1$ V, a dynamic range of $\text{DR}_u = (V_{\text{DD}}/2)^2/v_{\text{nu}}^2 \approx 50$ dB.

If we assume that the signal is a random bitstream, we will have $P_r \approx 2$, and if we also assume that the temperature is $T = 290$ K, the average energy consumption per sample will be $E_u = \text{DR}_u \cdot (kT/2) \cdot P_r \approx 10^{50 \text{ dB}/10 \text{ dB}} \cdot (1.38 \times 10^{-23} \text{ J/K} \cdot 290 \text{ K} / 2) \cdot 2 = 10^5 \cdot (2.0 \times 10^{-21} \text{ J}) \cdot 2 = 0.40$ fJ.

With a sampling rate of e.g. $f_s = 1$ GHz, we get $P_{\text{avg}} = E_u \cdot f_s = 0.40 \text{ fJ} \cdot 1 \text{ GHz} = 0.40 \mu\text{W}$. A complete system would of course consist of much more than a single minimum-size inverter and would therefore consume more power than this.

A different way to view the dynamic range is that we are $(V_{\text{DD}}/2)/v_{\text{nu}} = 250$ standard deviations away from accidentally crossing the $V_{\text{DD}}/2$ threshold when the signal is at 0 V or V_{DD} . In other words, the dynamic range is much higher than what a digital system actually needs. This prompts the question of why the devices have been designed this way. One possible explanation is that it could be a side effect of speed requirements. If the gate oxide was thicker, the capacitance would be lower and thus the power consumption lower too, but this would also lead to a higher channel resistance and thus a lower maximum sampling rate. A different explanation is that the high dynamic range might provide a margin which protects against errors caused by e.g. PVT variations and device mismatch.

It might be tempting to increase the gate lengths in an attempt to reduce the dynamic range, but this will not actually work. The inverter will become weaker in the sense that R_u increases, but this does not affect the dynamic range of sampled systems, as shown earlier. C_u will also

increase, though, and since this increases the dynamic range, the net effect of increasing the gate lengths in a minimum-size inverter will be that the dynamic range in fact increases.

A reduction of V_{DD} , on the other hand, will actually have an effect, so this is a viable method for reducing the unnecessarily high dynamic range and power consumption. The system will in practice be just as reliable with a reduced V_{DD} , but the energy consumption for a given computation will be lower than before. The drawback is that the maximum sampling frequency, i.e. the maximum clock frequency, will be lower too. To reduce V_{DD} like this is a well-known technique which is used in applications such as so-called subthreshold circuits.

Parallelization of unit-element amplifiers:

$$R = R_u/n$$

$$C = C_u n$$

Frequencies:

$$f_c = 1/2\pi R_u C_u = 1/2\pi RC$$

$$BW_{\text{noise}} = f_c \cdot \pi/2 = 1/4RC$$

$$f_{\text{su}} = 2 \cdot BW_{\text{noise}} = 1/2RC$$

$$T_{\text{su}} = 1/f_{\text{su}} = 2RC$$

$$OSR_u = f_{\text{su}}/f_{\text{s-Nyquist}} = 1/2RCf_{\text{s-Nyquist}}$$

$$OSR = f_s/f_{\text{s-Nyquist}}$$

Dynamic range:

$$v_{\text{nu}} = \sqrt{4kTR_u \cdot BW_{\text{noise}}} = v_{\text{nu}} = \sqrt{kT/C_u}$$

$$v_n = \sqrt{4kTR \cdot BW_{\text{noise}}} = v_{\text{nu}}/\sqrt{n} = \sqrt{kT/C}$$

$$v_{\text{n-Nyquist}} = \sqrt{4kTR \cdot BW_{\text{noise}}} / \sqrt{OSR} = v_{\text{nu}}/\sqrt{n \cdot OSR} = \sqrt{kT/C} / \sqrt{OSR}$$

$$DR_u = v_{\text{signal-max}}^2/v_{\text{nu}}^2 = DR_u = \frac{1}{2}C_u v_{\text{signal-max}}^2 / (kT/2)$$

$$DR = v_{\text{signal-max}}^2/v_n^2 = DR_u \cdot n = \frac{1}{2}C v_{\text{signal-max}}^2 / (kT/2)$$

$$DR_{\text{Nyquist}} = v_{\text{signal-max}}^2/v_{\text{n-Nyquist}}^2 = DR_u \cdot n \cdot OSR = \frac{1}{2}C v_{\text{signal-max}}^2 \cdot OSR / (kT/2)$$

Power consumption:

$$P_u = P_{\text{u-ideal}} \cdot P_r = (v_{\text{signal-max}}^2/4R_u) \cdot P_r$$

$$P = P_{\text{ideal}} \cdot P_r = (v_{\text{signal-max}}^2/4R) \cdot P_r$$

$$E_u = P_u \cdot T_{\text{su}} = \frac{1}{2}C_u v_{\text{signal-max}}^2 \cdot P_r = DR_u \cdot (kT/2) \cdot P_r$$

$$E = P \cdot T_{\text{su}} = \frac{1}{2}C v_{\text{signal-max}}^2 \cdot P_r = DR \cdot (kT/2) \cdot P_r$$

$$E_{\text{Nyquist}} = P \cdot T_{\text{su}} \cdot OSR = \frac{1}{2}C v_{\text{signal-max}}^2 \cdot OSR \cdot P_r = DR_{\text{Nyquist}} \cdot (kT/2) \cdot P_r$$

$$P_{\text{avg}} = E \cdot f_s = \frac{1}{2}C_u v_{\text{signal-max}}^2 \cdot n \cdot OSR \cdot f_{\text{s-Nyquist}} \cdot P_r = DR_{\text{Nyquist}} \cdot (kT/2) \cdot f_{\text{s-Nyquist}} \cdot P_r$$

Figure 3.4: Dynamic range and power consumption of the unit-element amplifiers, both with and without integration.

3.4.3 Analog, Multi-Bit, and 1-Bit Modulations/Encodings

In order for an electronic system to process or convey a data sample, e.g. a voltage value, it needs to expend energy. The higher the dynamic range requirement is, the higher the energy consumption will be. This trade-off was discussed in detail in section 3.4.1 and 3.4.2.

These native samples are not very energy efficient in certain cases, though. By utilizing modulations or encodings, the dynamic range requirements and thus the energy consumption can be drastically lowered, so in the current section, we will explore the energy consumption of systems that use such modulations and encodings. The analysis will be based on a simple model consisting of a modulation or encoding on top of native samples.

Modulations and encodings have benefits beyond what we find with this first-order analysis, though. Digital systems, for example, exhibit some of these benefits since all such systems use modulations/encodings rather than just native samples. For one thing, when compared to analog systems, digital systems provide much better noise immunity. This might come at the cost of a higher power consumption, but it could still be a desirable trade-off to make. For another thing, digital systems have tunable sampling rate. This could be used to reduce power consumption compared to analog systems since the latter operate continuously and can therefore not be duty cycled like a digital system can.

In order to find out how much energy a system consumes to convey one of these native samples, we will use a system model based on a unit-element amplifier. In the current section, we will use a simplified model of the unit element, though, not concerning ourselves with the details found in the two previous sections. The detailed unit element analysis in the previous sections and the modulation/encoding results from the current section could be combined in order to e.g. find actual energy consumption figures for a system that utilizes a given modulation. We will not perform this combination, though, as it should be quite straightforward.

3.4.3.1 Unit-Element Amplifier

We will first take a closer look at the trade-off between the dynamic range of a system and its power consumption. The dynamic range is the ratio between the largest undistorted signal and the noise of the system, or in other words the highest SNR value a signal could have in such a system. The noise can often be reduced at the cost of increasing the power consumption of a system, and this gives us the ability to trade off between a large dynamic range and a low power consumption.

To help us analyze the systems, we will establish a unit-element amplifier which can be used to transport or process a signal. The unit element is finite both in transistor gate width and in the time dimension, i.e. it works with sampled data, just like a switched-capacitor circuit does. We assume that the unit element has a dynamic range of 0 dB, i.e. that the maximum signal power and average noise power in a single sample have the same value. This means the gate width must be small enough and the time interval short enough to give that much noise. The reason the time interval has anything to say, is that if it is increased, the noise in a sample will be averaged over time and thus reduced. A dynamic range as low as 0 dB might not be achievable in practice, but this unit element is still a useful tool for the analysis.

We further assume that the noise is white and Gaussian, i.e. that two subsequent samples

have uncorrelated noise. We do not make any assumptions about the distribution of the signal, though, such as that it is Gaussian, uniform, sinusoidal, or binary-valued. The reason for this is that we will be working with rough estimates here, so that the difference between distributions like these will not have much to say.

Integration In order to convey a signal with a high SNR by using these unit elements, we will have to do some kind of integration. One way of doing this is to average multiple parallel unit elements, or in other words to use larger gate widths. Another way is to integrate over time. We do not need to do actual integration over time, though. Instead, we can just oversample by e.g. a factor of 100. The noise power will then be spread over a bandwidth which is 100 times larger, and since the noise amplitude per unit element sample stays the same, the total power will stay the same too, so that the PSD will be reduced by a factor of 100, i.e. 20 dB. We can then just ignore the high-frequency noise and only filter it out if we need to. The oversampling in this example will therefore increase the in-band dynamic range for a band-limited signal by 20 dB.

An analog system can be said to be such a sampled-data system where we oversample as much as possible, i.e. so that there are no gaps between the unit element time intervals. A sampled digital system, on the other hand, can be said to be such a sampled-data system where we integrate by parallelization, i.e. by using larger gate widths, and not by oversampling. For each clock cycle, the digital signal will then be conveyed with sufficient SNR to avoid bit errors, and this only takes a single time unit.

Energy Consumption The integration discussed here increases both the power consumption and the dynamic range of a system. More specifically, the power consumption is proportional to both the sampling rate of the unit elements and the parallelization factor, and the same goes for the dynamic range when measured as a power ratio. This means that the energy E_{Nyquist} needed to convey one sample of the original Nyquist-rate signal, i.e. not an oversampled sample, is $E_{\text{Nyquist}} = E_u \cdot 10^{\text{DR}_{\text{dB}}/10 \text{ dB}}$, where E_u is the energy consumed by one unit element during one unit of time and DR_{dB} is the desired dynamic range in decibels. The equation is useful in assessing different system configurations and will help us in the analysis below. In addition to E_u and E_{Nyquist} , we will also use the quantity E , which is the energy consumed by a parallelized unit element during one unit of time, i.e. E_{Nyquist} without the oversampling.

3.4.3.2 High-SNR Signals

We will here investigate how various modulations/encodings perform in terms of power consumption when the signal of interest has a high SNR. To exemplify this case, we will use a typical 16-bit audio signal. Such a signal can have an SNR of up to approximately 100 dB, so the system that we send the signal through gets a dynamic range requirement which is also about 100 dB.

Analog In order for an analog or sampled-data system to achieve this dynamic range, its energy consumption would have to be about $E_{\text{Nyquist}} = E_u \cdot 10^{100 \text{ dB}/10 \text{ dB}} = 10^{10} \cdot E_u$ per sample.

Multi-Bit A classical digital system, on the other hand, uses 16 bit per sample to represent the signal and would therefore need to handle 16 separate signals, each with an SNR high enough to avoid bit errors. If the noise is so low that the logic value threshold $V_{DD}/2$ is e.g. 7σ away from both logical 1 and logical 0, the Bit Error Rate (BER) will be 10^{-12} . The dynamic range required for this is about 17 dB, which we will round to 20 dB. This gives an energy consumption per data bit of about $E = E_u \cdot 10^{20 \text{ dB}/10 \text{ dB}} = 10^2 \cdot E_u$ and a total energy consumption per sample of about $E_{\text{Nyquist}} = 16 \cdot 10^2 \cdot E_u$.

1-Bit w/ SSR Although the in-band SNR of an SSR system can be high, each of the 1-bit samples will have a low SNR and are therefore very similar to the unit elements with 0 dB dynamic range. Both analog systems and SSR systems achieve better performance than these basic elements by integrating. This means that an SSR system will require about the same amount of integration as an analog system in order to achieve the same in-band dynamic range. Each 1-bit sample has an energy consumption of about $E = 10^2 \cdot E_u$, though, so the total energy consumption will be about $E_{\text{Nyquist}} = 10^{10} \cdot 10^2 \cdot E_u = 10^{12} \cdot E_u$ per Nyquist-rate sample of the signal, i.e. about 100 times the energy consumption of analog systems.

1-Bit w/ SDM SDM systems, on the other hand, are more efficient than SSR systems. As an example, the Direct Stream Digital (DSD) format of Super Audio CD (SACD) is a bitstream with an oversampling factor of 64 and thus with a data rate which is four times higher than normal 16-bit audio. An SDM system based on this format will therefore have an energy consumption of about $E_{\text{Nyquist}} = 64 \cdot 10^2 \cdot E_u$ per sample, i.e. almost the same as for a classical digital system. This means that SDM systems have a much lower power consumption than both analog systems and SSR systems, at least for the high-SNR case.

Discussion As we can see, there is a drastic difference in power consumption between analog and classical digital systems in this case. We might not see quite as large a difference in practice, though, since digital circuits are often optimized for speed. Among other things, this leads to a higher supply voltage, which leads to a higher dynamic range and thus a higher energy consumption E per bit. This unnecessarily high dynamic range is discussed further in section 3.4.2.10. The subthreshold design paradigm, with its low supply voltage, would probably have an energy consumption closer to what we have found here, though.

An important lesson from this comparison of modulations/encodings, is that having a high dynamic range is much more costly than having multiple signals. This is something we can also see in the Shannon-Hartley channel capacity theorem [W], $C = B \cdot \log_2(1 + \text{SNR})$, where a doubling of the SNR, i.e. a doubling of the power consumption, only gives one more bit of channel capacity for $\text{SNR} \gg 1$, whereas a doubling of the bandwidth B , i.e. also here a doubling of the power consumption, would in fact double the number of bits. It is therefore most power efficient to keep the SNR low and rather use parallel radio channels, wires, or similar when trying to increase the information throughput. Note that in our case, a quadrupling of the SNR, i.e. a doubling of the signal amplitude, is needed in order to get one more bit of resolution in the samples. This is because our samples are not complex-valued like radio channels are.

3.4.3.3 Low-SNR Signals

As we will show here, the results are quite different for low-SNR signals than for high-SNR signals. To exemplify the low-SNR case, we will use a signal with an SNR of 0 dB. SNR values as low as or lower than this can be found e.g. where spectrum spreading techniques such as CDMA are used, for example in GPS receivers.

Analog Analog and sampled-data systems will have an energy consumption of about $E_{\text{Nyquist}} = E_u \cdot 10^{0 \text{ dB}/10 \text{ dB}} = 1 \cdot E_u$ per sample. In other words, such systems would use a single unit element per sample, i.e. they would not perform any form of integration.

Multi-Bit Classical digital systems could either be said to degenerate to the 1-bit SSR system here or they could still be used with multi-bit resolution. One benefit of the latter option is that the signal and noise will stay inside the ADC range so that the system behaves as normal rather than having the out-of-range peculiarities of an SSR system. Another benefit is that the quantization error, which is basically just additional noise, can be arbitrarily reduced with an increase in bit depth. If we choose to use multi-bit resolution rather than 1-bit resolution, for example four bits, the energy consumption would be about $E_{\text{Nyquist}} = 4 \cdot 10^2 \cdot E_u$ per sample.

1-Bit w/ SSR SSR systems could just 1-bit quantize the sampled-data samples without needing to do any form of integration such as oversampling. This approach would not cause much SNR loss since the information content per sample already is about 1 bit. SSR systems will therefore have an energy consumption of about $E_{\text{Nyquist}} = 10^2 \cdot E_u$ per sample, i.e. almost the same as for classical digital systems. For any SNR value less than about 0 dB, an integration factor of just $\pi/2 \approx 1.57$ is enough to compensate for the SNR loss caused by the 1-bit quantization. This would give $E_{\text{Nyquist}} = 1.57 \cdot 10^2 \cdot E_u$. Exactly how this integration must be performed, can be learned from chapter 4.

1-Bit w/ SDM SSR systems are good enough with little or no integration here. SDM systems, on the other hand, require some oversampling and are also a bit more complex than SSR systems. This means that SDM systems will not provide huge power savings compared to SSR systems in this case, like we saw them do for high-SNR signals.

Discussion As expected, analog systems and SSR systems, both of which use a simple form of integration, are much more power efficient for this low-SNR case than for the high-SNR case. This is because a high dynamic range costs a lot of power when using such simple integration.

Furthermore, analog systems are here actually more power efficient than digital systems, both 1-bit and multi-bit, since the data bits require a 20 dB dynamic range, which is more than the 0 dB needed by analog systems in this case. In a practical system, however, we probably have a few decibels of dynamic range anyway since systems seldom have noise amplitudes on the order of V_{DD} . This means that analog samples and the data bit voltages will have similar dynamic ranges and thus similar energy consumption. So in practice, analog systems might not be better than digital systems in this case after all.

We also see that SSR systems are slightly more power efficient than classical digital systems. This is simply because they use fewer bits per Nyquist-rate sample.

3.4.3.4 Forced Oversampling in Analog Systems

An interesting observation to make about the low-SNR case, is that for signals with low bandwidth, analog systems integrate way too much because of their large oversampling of the unit elements. A practical sampled-data system will easily have enough dynamic range to handle the low-SNR signal without the use of any form of integration, i.e. just using minimum-size amplifiers with the sampling rate set to the Nyquist rate. Since analog systems are basically oversampled versions of this, they could in practice be using an oversampling factor on the order of millions or billions even though such integration is completely unnecessary. This could represent a large waste of power.

Since analog systems will always have this forced oversampling, we will have to find other ways to reduce their power consumption. This is not so easy, though. For digital systems, increasing the gate lengths does not help at all but decreasing the supply voltage does. This was shown in section 3.4.2.10. We can do the same kind of analysis for analog systems, also here using the terms and concepts from section 3.4.1 and 3.4.2. If we increase the gate lengths, both R and C will increase as a result. The consequence of this is that both the oversampling factor of analog systems, $OSR_u = 1/2RCf_{s-Nyquist}$, and the total power, $P = (v_{\text{signal-max}}^2/4R) \cdot P_r$, will be reduced. However, the energy consumption of each of the oversampled samples, $E = \frac{1}{2}Cv_{\text{signal-max}}^2 \cdot P_r$, will actually be increased. This means that when the gate lengths are so large that $OSR_u = 1$, the energy consumption of each sample will be much larger than for the minimum-size amplifier. Even though increasing the gate length reduces P , we can not do this arbitrarily much, since $OSR_u \geq 1$ sets a limit for how large the gate lengths can be. So, increasing the gate lengths helps, but only up to a certain point. A different approach would be to weaken the amplifiers in other ways, e.g. by starving them. This might perhaps solve the problem, but we will not look further into such solutions here. Yet another solution is to reduce the supply voltage. This would work here too, just as it did for digital systems in section 3.4.2.10. There might be a limit to how much we can reduce it in practice, though.

As explained earlier, SSR systems are very similar to sampled-data systems for the low-SNR case, so both SSR systems and sampled-data systems are good alternatives to analog systems. The reason for this is that they do not have the forced oversampling. Even for sampled systems, though, it is hard in practice to reach a low dynamic range and thus a low energy consumption, as shown in section 3.4.2.10. This means that SSR systems, which require a dynamic range of about 20 dB for the data bits, and sampled-data systems, which could have a requirement of maybe just 0 dB, would have similar performances if the dynamic range could not go lower than e.g. 20 dB. Therefore, SSR systems might actually in practice be just as good as sampled-data systems, both of which are better than analog systems. Due to the very large forced oversampling of analog systems, SSR systems might be able to perform some integration and still be more power efficient than the analog systems. Such SSR systems would be able to handle signals with moderate SNR levels.

3.4.3.5 Summary

Dynamic range costs power. This means that it takes a lot of power to process high-SNR signals by using native samples. A way to get around this problem, is to use modulations/encodings. These can e.g. translate the high dynamic range requirement of an original sample to a lower dynamic range requirement, at the cost of using more than just one sample. Different modulations/encodings have different properties and power consumption levels, which also depend on the SNR of the signal. This is summarized below, where a low signal bandwidth has been assumed:

- High SNR
 - Analog / SSR: High power
 - Multi-bit / SDM: Low power
- Low SNR
 - Analog (continuous-time): High power (due to forced oversampling)
 - Analog (sampled-data) / multi-bit / SSR / SDM: Low power
 - SSR uses slightly less power than multi-bit
- Low/moderate SNR
 - SSR might use less power than analog due to forced oversampling in analog systems

An interesting observation to make from this, is that the power consumption of simple 1-bit SSR systems can be similar to or even lower than the power consumption of analog or classical digital systems, at least as long as the SNR of the signal is not too high.

Note that the power consumption analysis that has been performed here is limited in scope since it is only based on considerations around the dynamic range of samples. The results found here might therefore not reflect practical systems precisely, but they do at least shed some light on a few theoretical aspects of power consumption. This also exposes several interesting opportunities for power saving.

3.4.4 Power Benefits of Digital Systems

In the previous sections, we have investigated power consumption with regards to dynamic range and modulations/encodings. There are other issues at play than just this, though, so in this section, we will look into a few other power consumption benefits we get from switching from analog to digital systems. Then, in the next section, we will look into the benefits of switching from multi-bit to 1-bit systems. When switching from analog to 1-bit systems, we therefore get both sets of benefits.

Analog systems are often used with back-off, i.e. with a reduced signal amplitude, in order to ensure linear behavior over the full signal range. This means that the dynamic range will be reduced without the power consumption being reduced too, which obviously represents an inefficiency. Digital systems, on the other hand, have their signals going rail to rail, so they do not suffer from this problem.

Analog systems might take some time to reach steady state after they are powered up. Digital systems, on the other hand, do not have this problem, and they may not even actually need to be powered down in order to enter power-saving mode. This means that digital systems can be duty cycled, i.e. be powered up just for short intervals, much more efficiently than analog systems.

Digital systems can use advanced algorithms that are more efficient than the straightforward solutions that analog systems might employ. Examples of this are data compression, FFT, and compressed sensing [W].

3.4.5 Power Benefits of 1-Bit Systems

We will here look into a few power consumption benefits we get from switching from multi-bit to 1-bit systems, that have not already been mentioned in the previous sections. All of these benefits are potential reasons for why e.g. SSR systems might be preferred over classical digital systems in certain cases.

1-bit systems have simpler circuits than multi-bit systems. The clock frequency might be higher, though. As an example, consider a low-SNR signal which needs to be multiplied with another signal, e.g. to perform cross-correlation. A 1-bit system could use a simple XNOR gate to perform the multiplication, whereas e.g. a 4-bit system would need a much more complicated multiplication circuit.

When using sigma-delta ADCs and DACs, decimation and interpolation circuits are required to convert to and from multi-bit signals. 1-bit systems could perform signal processing on the bitstreams directly, thus eliminating those circuits and their power consumption.

Continuous-time 1-bit systems, i.e. CTBV systems, have additional benefits. Since there is no constantly running clock in these systems, they will only consume power when the signals change value. This might represent a huge power saving for applications where events are far apart.

CTBV systems can also have a very quick response time for handling events. This comes at no additional cost to these systems. Classical digital systems, on the other hand, would have required high clock frequencies with high power consumption in order to achieve the same performance.

3.5 Example Applications

1-bit signal processing can be used for many different applications. We will here list a few.

SSR or similar techniques, i.e. thresholding of noisy samples, can be found in e.g. sonar applications, CDMA systems such as GPS receivers, M-sequence radar, photographic film, and evolution with its life/death 1-bit quantization.

A different type of sampled 1-bit systems are those based on SDM techniques. These share some properties with SSR systems but are more complex.

There are also continuous-time 1-bit systems. In the following examples, the signals consist of events that are single points in time, i.e. the events have no duration or the duration does not contain any information: Photons, tunneling or otherwise jumping electrons, and nerve signals in neurons.

Continuous-time systems can also have events with an associated duration. This can be seen in e.g. the half-tone printing process, class-D amplifiers with their use of PWM, PWM control of LED lights, and PWM control signals for servos.

1-bit signal processing is suitable for the combination of WSN, UWB, and CMOS due to the low SNR, simplicity, and low power consumption.

3.6 1-Bit Signal Processing As a Fundamental Principle

1-bit signal processing is an important tool in technology. It is utilized in SDM systems, which are in widespread use, and in SSR systems, which could be said to be digital systems in their simplest and most basic form.

We can also observe 1-bit principles at play in natural processes, e.g. in photographic film, neurons, and evolution.

There might be something even deeper than this about the 1-bit concept, though. Light can be said to be a 1-bit signal since discrete photon events carry the analog luminosity signal. Also electrons that tunnel or otherwise jump, such as in vacuum tubes, are discrete events like this and therefore a type of 1-bit signals. The reason for this is simply the laws of nature. Elementary particles like photons and electrons are discrete quanta that do not carry any amplitude information themselves. In other words, a stream of particles like this is a continuous-time 1-bit signal.

These observations make the tantalizing suggestion that there is something fundamental about 1-bit signal modulation. An evocative way to put it, is like this: There is a reason why 1-bit quantization and quantum physics are similar words.

Chapter 4

Sampled 1-Bit Signal Processing

Contents

4.1	Introduction	65
4.2	Other Uses for 1-Bit Sequences Than Conveying Signals	66
4.3	Suprathreshold Stochastic Resonance	67
4.3.1	Stochastic Resonance	67
4.3.1.1	Observations and Applications of Stochastic Resonance	69
4.3.2	Subthreshold Stochastic Resonance	69
4.3.3	Suprathreshold Stochastic Resonance	69
4.3.4	Trading off SNR and Sampling Rate	73
4.3.5	Effects of Noise, Nonlinearity, and Integration	73
4.4	Comparison with Sigma-Delta Modulation	77
4.5	Threshold Levels	78
4.5.1	Swept-Threshold Sampling	78
4.6	Integration by Oversampling or by Parallelization	80
4.7	General Benefits of Oversampling	81
4.7.1	Oversampling in the Retina	82
4.8	Decoding	84
4.9	SNR Metrics	85
4.9.1	Measurement Setup	85
4.9.2	Gains	86
4.9.3	SNR Metric A: Consider Error in Absolute Value	88
4.9.4	SNR Metric B: Maximize SNR Value	89
4.9.5	SNR Metric C: Correctly Identify Signal Part	91
4.9.6	Examples of the Differences	92
4.9.7	Oversampling	92
4.9.8	Distortion Compensation	93
4.9.9	Limitations of the SNR Metric	93
4.10	Small-Signal Analysis	94
4.11	Large-Signal Analysis	100
4.11.1	With an Input Signal Distribution	100
4.11.2	Integration	101
4.11.3	Uniform Noise	101
4.12	SNR As a Function of Input Noise	105
4.12.1	SNR Metrics A, B, C	106
4.12.2	Quantizer Output Levels	106
4.12.3	Related Work	110
4.13	Colored Noise	112
4.13.1	Why SSR with Colored Noise Works	113
4.13.2	White Noise from First-Order Anti-Aliasing Filter	114
4.14	Spectrum Simulations	115
4.14.1	Measuring the SNR	115
4.14.2	Overview	117
4.14.3	Simulations	117
4.15	Bitstream Processing	134
4.16	Example Applications	135

4.1 Introduction

This chapter is largely about the effects of 1-bit quantization of signals with noise, or in other words, about converting analog signals to bitstreams. We will only consider straightforward 1-bit quantization, known as Suprathreshold Stochastic Resonance (SSR), not the more complex Sigma-Delta Modulation (SDM) technique. We will also limit ourselves to sampled systems. Continuous-time systems, on the other hand, will be covered in chapter 5.

Note that we will be using the terms “SSR” and “1-bit quantization” interchangeably even though SSR strictly speaking only applies when the noise is beneficial. If the end-result SNR after all processing is desired to be higher than a few decibels, which often is the case, there will probably be beneficial noise involved, though. It is therefore usually completely correct to use the SSR term. The beneficial noise is an effect that will be explored in detail in this chapter.

Converting Analog Signals to 1-Bit Signals An analog signal can be converted to a sampled 1-bit signal without losing much information. If we let the 1-bit signal have a high density of “1”s when the analog signal is high and a low density when it is low, then by low-pass filtering the 1-bit signal later, we get back an analog signal close to the original signal. This is known as Pulse-Density Modulation (PDM). In other words, the 1-bit signal contains the input signal in the low-frequency part of its spectrum, while the high-frequency part of its spectrum just contains noise. The higher the sampling rate of the 1-bit signal is, the more precise the density values will be and thus the better the SNR of the signal of interest will be. In the frequency domain, the higher sampling rate means that the noise can be spread over a wider frequency band, thus reducing the amount of noise that falls in the signal band. Several different methods can be used to produce the 1-bit signal, such as SSR, SDM, and PWM.

SSR Compared to Other Methods A PDM signal can be made with e.g. SSR, SDM, or PWM. These methods differ in system complexity and might be better or worse solutions depending on e.g. the characteristics of the input signal and the input noise. For example, SSR requires the presence of input noise and if that noise has to be added explicitly, it reduces the SNR of the input signal, which other methods such as SDM would avoid. However, if there is already enough noise in the input, an SSR system, perhaps with filtering before quantization, i.e. with colored noise, would be a good and simple solution. In such a case, there might be very little loss caused by choosing the simpler SSR solution over more complex solutions like SDM.

4.2 Other Uses for 1-Bit Sequences Than Conveying Signals

1-bit sequences can also be used for other purposes than conveying analog signals, for example different forms of data bit serialization.

It is used in serial links such as SPI, USB, SATA, and PCIe for reasons such as that there will be no timing skew between parallel bits and that there will be fewer wires.

It can also be used to make 1-bit CPUs [Maye 02 ☹]. This gives smaller circuits at the cost of slower operation.

4.3 Suprathreshold Stochastic Resonance

We will here introduce Suprathreshold Stochastic Resonance (SSR), which is the focus of the current chapter. This is a specific type of the Stochastic Resonance (SR) phenomenon, so we will also discuss that here.

4.3.1 Stochastic Resonance

Stochastic Resonance (SR) is a phenomenon where noise improves the performance of a system, or in other words “beneficial noise”. This might at first sound counterintuitive, but for nonlinear systems, this is a very real effect. SR is basically the same as dithering, but in a wider sense.

Figure 4.1 shows the typical relationship between system performance and input noise strength. If there is no noise, the performance will be bad. If there is some noise, the signal gets through the nonlinear elements much better, and the performance is good. If there is too much noise, the performance will be bad again. In other words, there is an optimal noise level, i.e. an SR peak. The performance metric could be e.g. SNR or some kind of information theoretic measure. Note that the position of the SR peak, i.e. the optimal noise strength, depends on the choice of metric.

The curve in the figure also shows why it is called “stochastic resonance”. The SR peak simply looks like a resonance peak [McDo 09b Ⓢ], and it is the strength of the noise, which we can call a stochastic signal, which is the x-axis of the plot. In dynamical SR systems, there is actually a form of frequency matching [McDo 09b Ⓢ], so the “resonance” term is quite appropriate there. The literal meaning of the name is not so important, though, since SR has now been established as a term for beneficial noise in general.

Figure 4.2 shows the SR effect in an image. After thresholding the original image, it loses a lot of information and is no longer recognizable. By adding noise before the thresholding, the thresholded image becomes much better, i.e. the noise was beneficial.

Integration can help improve the performance of SR systems, and for some systems, this is actually necessary in order for the SR effect to occur.

So to sum up, an SR system consists of noise, nonlinearity, and maybe integration. This means that we can use circuit elements that are simple, cheap, and bad, i.e. noisy and nonlinear, and just integrate many outputs from them. This integration can be done either with a high clock frequency or with parallelization, both of which are feasible solutions due to the simplicity of the elements. This paradigm therefore gives us good system performance even though we only use crude circuits.

There are many aspects of SR systems that can be varied. A few of them are as follows:

- Dynamical / non-dynamical
- Periodic / aperiodic
- Subthreshold / suprathreshold
- 1-bit quantizer / multi-level quantizer / just soft nonlinearity / clipping
- Same / different threshold levels
- One / many nonlinear circuits
- Parallel hardware / integration in time / no integration

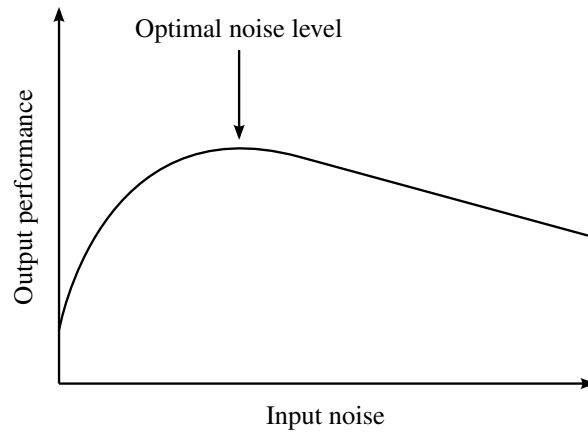


Figure 4.1: Adding moderate amounts of noise to the input can actually improve the output performance of some nonlinear systems. This phenomenon is called Stochastic Resonance (SR).

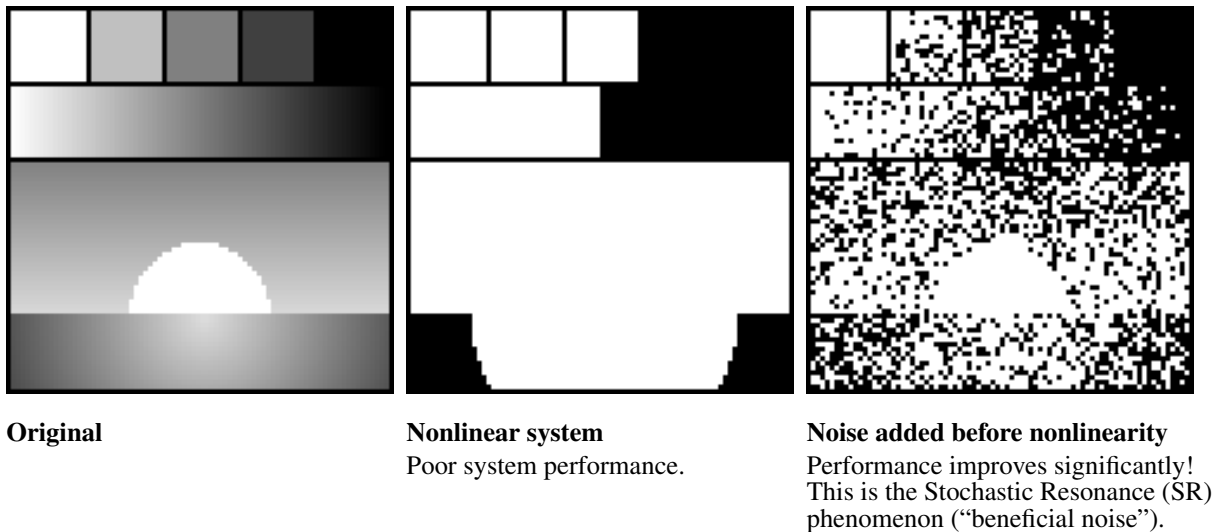


Figure 4.2: SR in an image.

- Same / different noise source for each element
- Coupled array (“array enhanced SR”)
- Which signal Probability Density Function (PDF)
- Which noise PDF
- White / colored noise

A lot of studies have been done on SR in all its forms and contexts. A few key references is the original SR paper [Benz 81 ⑤], the associated paper connecting SR to climate [Benz 82 ⑤], an overview article [Rouv 07 ⑤], and a PhD thesis covering the subject [McDo 06b ⑤]. SR was originally identified in dynamical systems, specifically climate models. We will not work with these types of systems here, though. Instead, we will only work with non-dynamical systems.

4.3.1.1 Observations and Applications of Stochastic Resonance

SR has been observed in chemical reactions, semiconductor devices, nonlinear optical systems, magnetic systems, and Superconducting QUantum Interference Device (SQUID) [Rouv 07 ⑤]. Furthermore, noise improves the sensory input in organisms and is also beneficial in neuron models [Wiki 12 ⑤]. This does not prove that SR is being used by neurons, though [McDo 09b ⑤].

SR has been used in applications such as vibrating insoles for improved balance in elderly people [Prip 03 ⑤] and cochlear implants [McDo 09a ⑤]. Dithering, which is basically the same as SR and which predates SR theory, is widely in use, e.g. in digital audio, images, and other applications where ADCs/DACs or similar mechanisms are in use. SR is also used in the proposed “stochastic electronics”, which are circuits that exploit noise [Hami 14 ⑤].

We could also look at SR and noise in a much wider sense. This allows us to identify SR at play both in cosmology and evolution. Inhomogeneities in the early universe caused matter to clump together to form galaxies, etc. Without this “noise”, the universe would still consist of completely homogeneous matter without any structure. Genetic mutations could be said to be “noise” that is beneficial to evolution. This observation suggests the possibility that organisms may have adopted to mutate at the optimal intensity, i.e. at the SR peak, since this would make the organisms evolve quicker.

4.3.2 Subthreshold Stochastic Resonance

Subthreshold stochastic resonance, which it should be noted is not abbreviated SSR, is when the signal is below the threshold of a nonlinear system. Noise is then required in order for any signal at all to pass through the system.

Figure 4.3 shows a block diagram of a general stochastic resonance system, which could be a subthreshold stochastic resonance system with the right threshold setting, and figure 4.7 shows the signals in such a system, particularly demonstrating the benefits of noise.

Note that integration is not required here, unlike for some other systems.

4.3.3 Suprathreshold Stochastic Resonance

Suprathreshold Stochastic Resonance (SSR) is when the signal is above the threshold of a nonlinear system. There will then be some information transfer even when there is no noise, but noise could help enhance the system performance.

Figure 4.4 shows a block diagram of an SSR system, and figure 4.8 shows the signals in such a system, particularly demonstrating the benefits of noise. Figure 4.5 shows a second block diagram, this time with parallel nonlinear elements. The two block diagrams represent two different ways to perform integration, namely oversampling and parallelization.

Figure 4.6 gives a few more details about the configuration and properties of a typical SSR system. The different properties of the noise can have much to say, such as the distribution, including its DC value, standard deviation, PDF shape, and PDF symmetry, and also the color of the noise. Additive White Gaussian Noise (AWGN) is often assumed. If the input is analog, the system will have to both sample and quantize the signal. The order of these two operations is interchangeable. SSR systems often require integration. Without this, noise might be detrimental

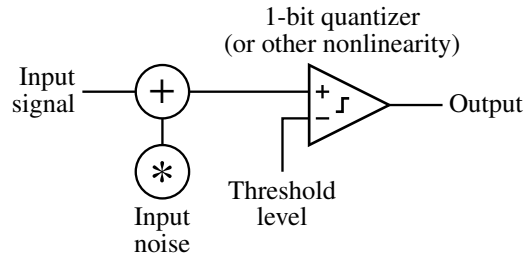


Figure 4.3: Stochastic Resonance (SR) block diagram. Optional signal integration not shown.

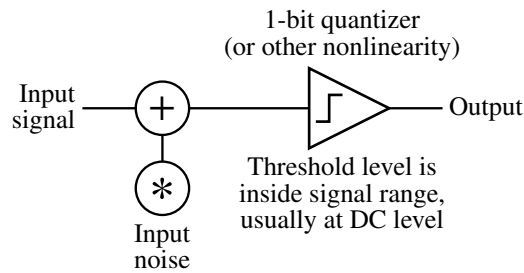


Figure 4.4: Suprathreshold Stochastic Resonance (SSR) block diagram. Integration of the output can be done over time, i.e. with low-pass filtering. The system is usually sampled.

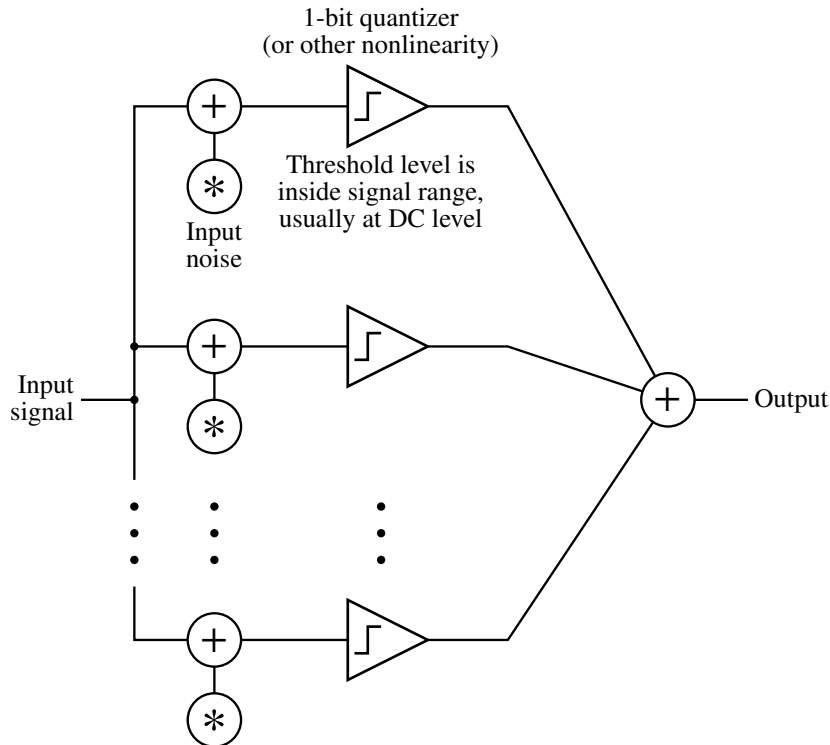


Figure 4.5: Suprathreshold Stochastic Resonance (SSR) block diagram with integration by parallelization.

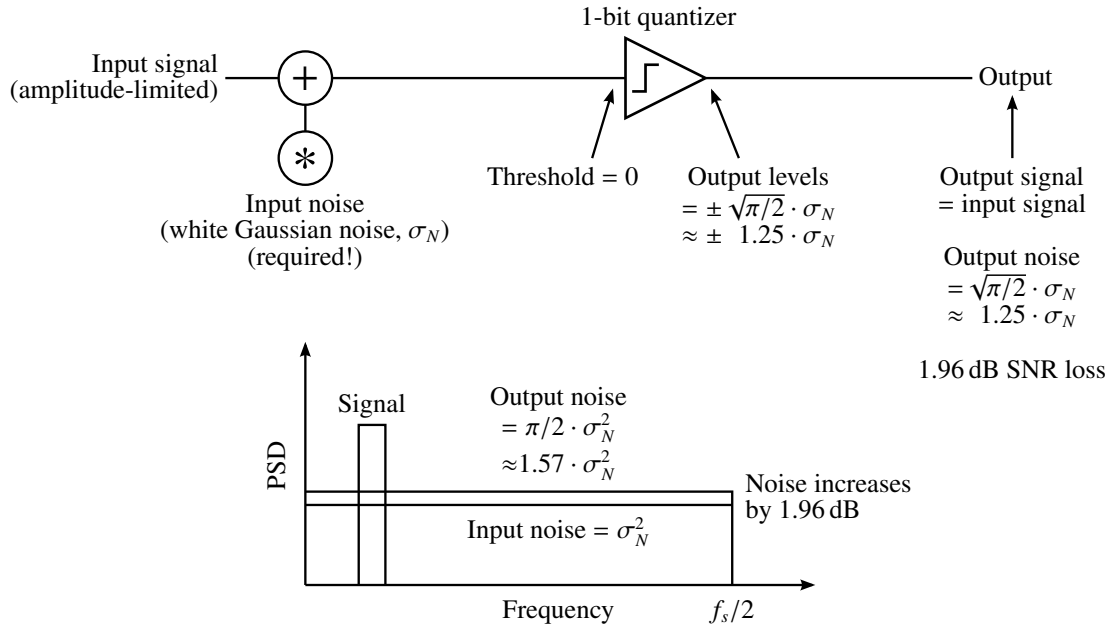


Figure 4.6: A typical sampled SSR system.

rather than beneficial. Integration is often done anyway, though. Even just a low-pass, band-pass, or other type of filter is integration, and so can other signal processing operations be too. An important result about SSR systems, is that the SNR loss due to the 1-bit quantization is only 1.96 dB when the input SNR per sample is low and the noise is white and Gaussian. Note that this result is not specific to SR, i.e. when the noise is beneficial. It also applies to 1-bit quantization in general, e.g. for very noisy inputs.

Note that the bitstreams from oversampled SSR systems are normal signals, i.e. the signal of interest is not modulated in a special way, there is just a lot of out-of-band noise. This is illustrated in the spectrum in figure 4.6. Another thing to note, is that SSR is not suitable for signals with a very high SNR, such as audio, since this would require huge amounts of integration.

We can see some similarities with other types of systems. For one thing, SSR is essentially a classical 1-bit ADC with dithering. A crucial difference is that the input is often outside the ADC range, though. This causes SSR and ADCs to behave differently. For another thing, both dithering ADCs, i.e. multi-bit without noise shaping, and SDM, i.e. 1-bit with noise shaping, are widely used techniques, whereas SSR, i.e. 1-bit without noise shaping, is less used because of its crudeness. Due to its simplicity, SSR might still have some interesting applications, though. Finally, the PWM technique can be seen as an SSR system with sawtooth noise. What happens in such a system, is basically that the threshold is swept. This is therefore very similar to what we call the swept threshold technique.

It can be interesting to ask what it is that makes dithering so important. If we look at a single sample in isolation, dithering only makes it noisier, but when we somehow integrate several samples, dithering is very important because it ensures that the expected values of the samples are correct. When we then integrate more and more, the result will get closer and closer to the correct value. Without dithering, the expected values will be wrong, and even infinite integration

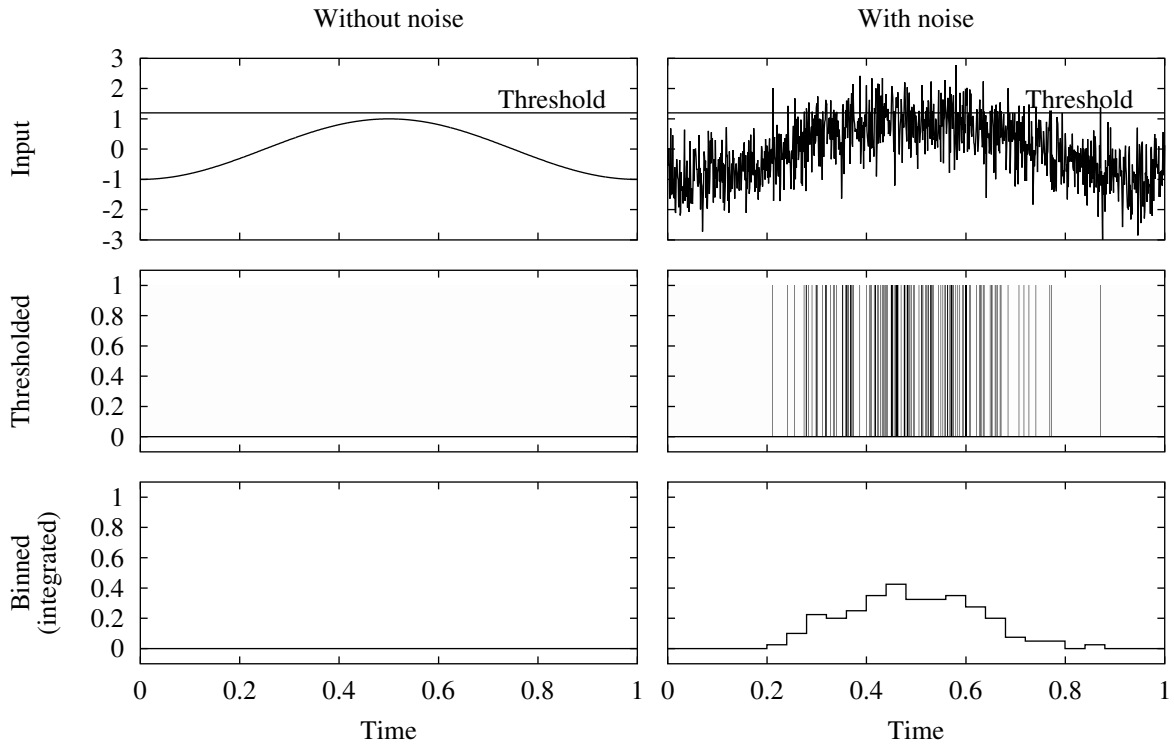


Figure 4.7: Demonstration of the benefit of noise in subthreshold SR.

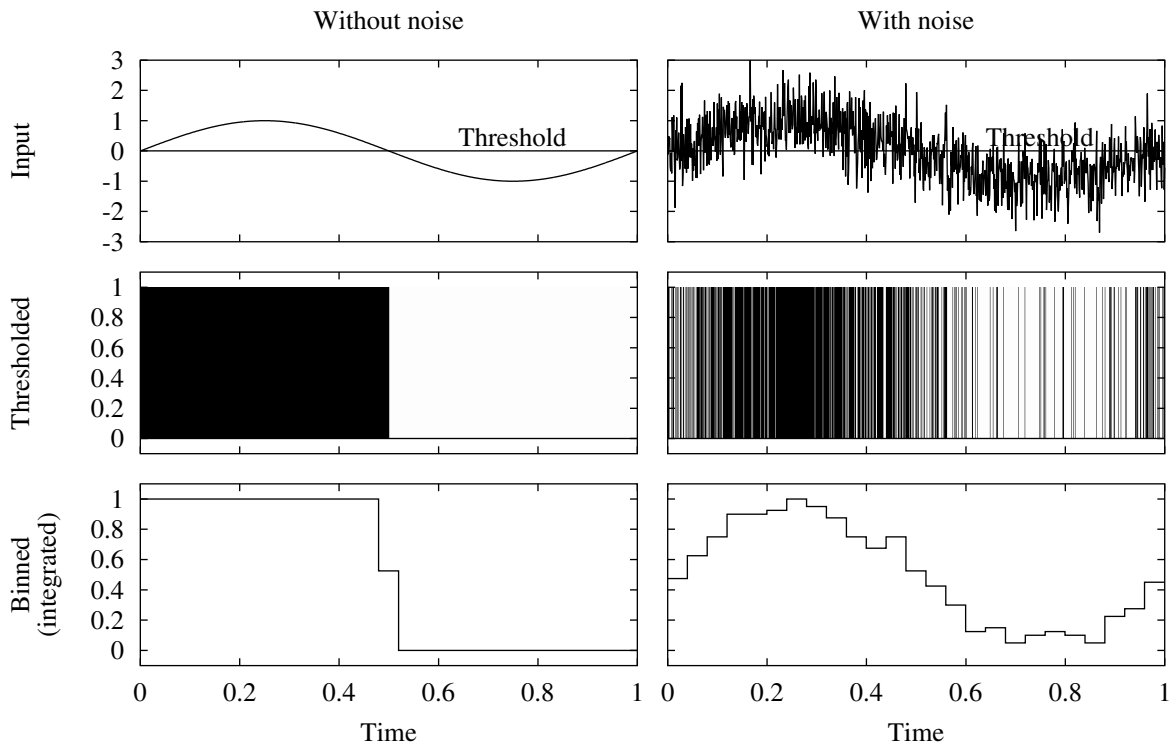


Figure 4.8: Demonstration of the benefit of noise in suprathreshold SR, i.e. SSR.

will then yield incorrect results. Dithering can be done in many different ways. We will here list a few methods in order of increasing complexity and statefulness in order to illustrate this tradeoff parameter: White noise, colored noise such as first-order low-pass filtered, deterministic “noise” for PWM mode, and error diffusion such as SDM.

The current chapter, chapter 4, explores many of the details concerning SSR systems.

4.3.4 Trading off SNR and Sampling Rate

In order for SSR to work properly, the SNR per sample needs to be low. Otherwise, the 1-bit quantization will destroy too much information content, or to put it another way, there will be too little dithering noise to ensure that the 1-bit samples get the correct expectation values. Fortunately, we can easily control this SNR value.

If we double the sampling rate f_s of a signal without changing the signal of interest but with an extension of the noise PSD into higher frequencies, the noise power per sample will double too. This means that the SNR per sample will be reduced by a factor of two. Note that only the full-band SNR changes. The in-band SNR, i.e. the SNR of the signal of interest, is unaffected. We then get that $\text{SNR}_{\text{sample}} \cdot f_s$ is constant. This degree of freedom allows us to choose what the SNR per sample should be.

By choosing an SNR per sample slightly below 0 dB this way, 1-bit quantization will not cause much information loss since the information content then already will be very low. To put it another way, the 1-bit samples will have correct expectation values due to sufficient dithering noise and we will be integrating a lot, so we will therefore be able to recover the signal properly. With the use of this technique, we can thus easily convert an analog signal to a bitstream without much SNR loss.

Figure 4.9 illustrates the tradeoff and the 1-bit quantization, both with overview plots and with noisy images as an example.

The image examples are quite close to what one might see in frames from a movie. When viewing a movie in SDTV resolution, there will be little noise in the frames, but when viewing it in HDTV resolution, the frames might be more noisy. What is most interesting here, though, is that the original photographic film, which has an even higher resolution than HDTV, is in fact binary, as discussed in section 4.16. The reason for this is simply that the silver-halide crystals in the film either get exposed or not, i.e. they have a binary response to light.

Note that we could either use inherent noise in the input or explicitly add noise ourselves. Also note that an important point here is that we can perform signal processing using simple 1-bit circuits instead of analog or multi-bit circuits. This might be cheaper and might even provide more advanced processing operations than an analog system could. The cost of this 1-bit alternative is that the sampling rate might be higher than for a multi-bit system.

4.3.5 Effects of Noise, Nonlinearity, and Integration

Noise, nonlinearity, and integration are some of the most important properties related to SR systems, but different systems, SR or not, may or may not have each of these properties. We will here compare the different system categories that result from these combinations.

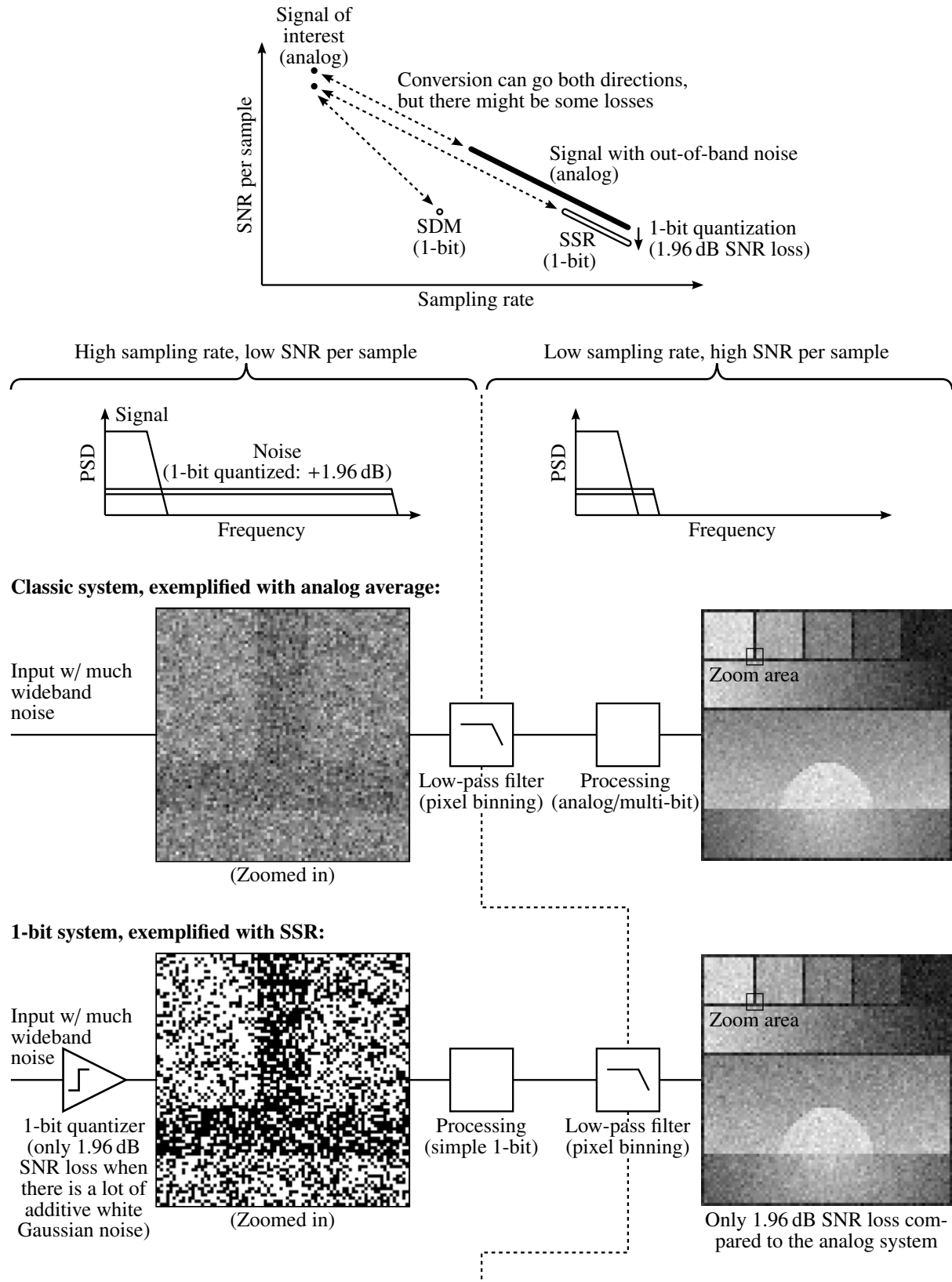


Figure 4.9: Top: A plot showing how SNR per sample and sampling rate can be traded off against each other without reducing the signal quality. Bottom: Examples of this in a classic system and a 1-bit system. When the SNR per sample is low enough, 1-bit quantization will not cause much loss of information.

Figure 4.10 lists reasons for why converting systems from different categories to the SR category might be beneficial. In other words, why adding noise and nonlinearity might be a good thing. This can help us choose when to use SSR for a system design.

Figure 4.11 lists the effects of adding noise and/or integration to a nonlinear system that has neither, both for subthreshold SR and SSR systems. The figure speaks for itself, but an important point is that SSR requires integration if we add noise. Without integration, the noise will just be detrimental, since it will then disturb the 1 bit of polarity information which we would otherwise have about the signal. Subthreshold SR systems, on the other hand, do not require integration in order to benefit from noise, since noise helps there be any threshold crossings at all. With properly chosen noise and integration, we can achieve arbitrarily good performance. Note that we here consider integration by parallelization rather than by oversampling. There are some small differences between the two. Also note that only the system configurations where there is beneficial noise are strictly speaking SR systems.

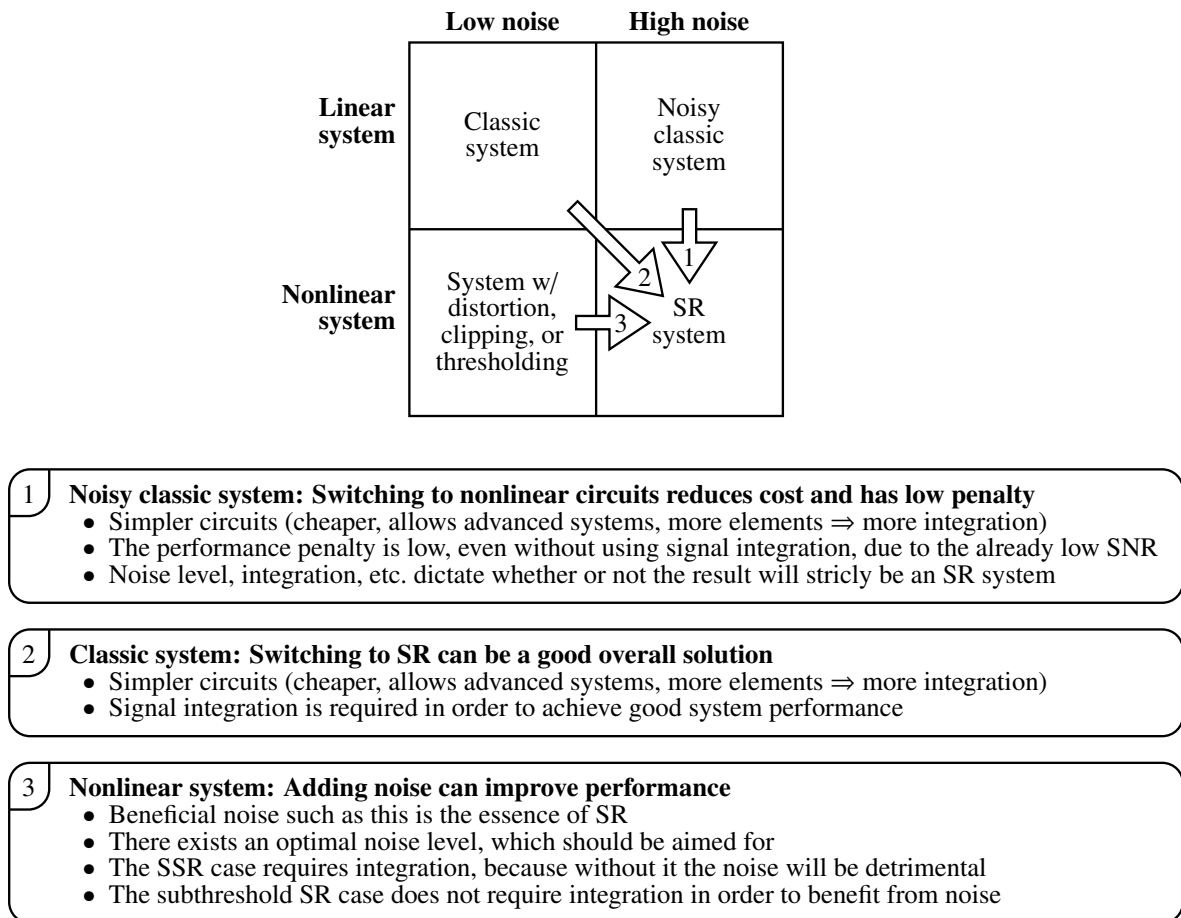


Figure 4.10: Different system categories and why moving to the SR category can be beneficial.

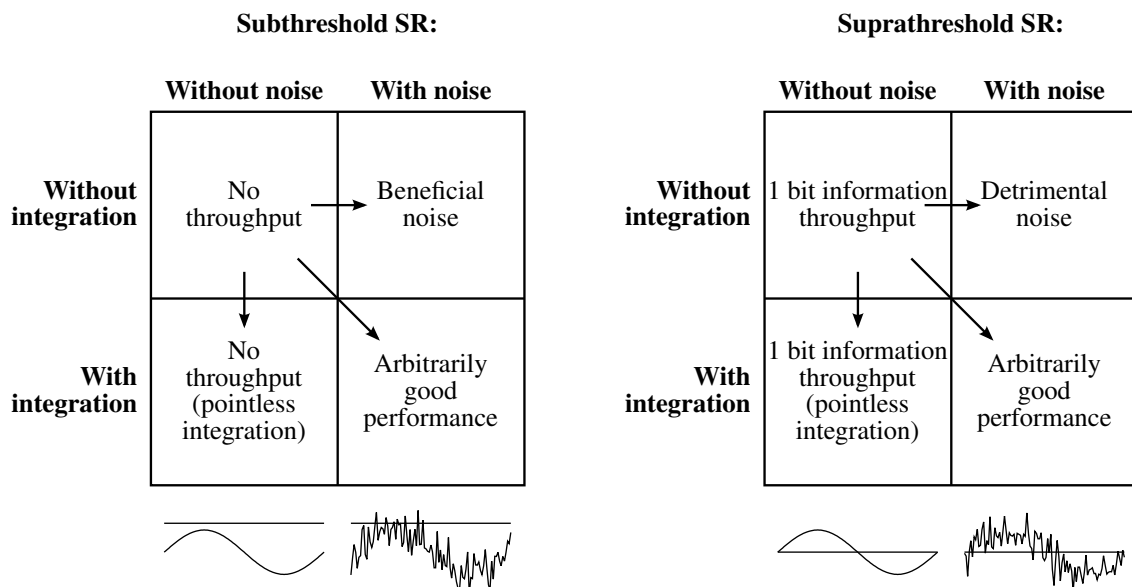


Figure 4.11: The effects of adding noise and/or integration to a nonlinear system that has neither.

4.4 Comparison with Sigma-Delta Modulation

An alternative to SSR is Sigma-Delta Modulation (SDM), also known as noise shaping or error diffusion. SSR, on the other hand, has unshaped noise. The difference between these two methods is illustrated in figure 4.12. As we can see, noise shaping greatly improves the in-band SNR. Note that we are here talking about oversampled SSR, not parallelized, and 1-bit SDM, not multi-bit, so both SSR and SDM are therefore here simple PDM bitstreams. The disadvantage of SDM systems is that they are more complex than SSR systems, so the higher SNR comes at a price.

If we want to convert a noiseless signal to a bitstream, we will have to add some form of dithering noise in order to achieve linearity. SDM shapes this necessary noise, or in other words adds it pre-shaped. Since it is the quantization noise of the converter itself which gets shaped, SDM will, of course, not be able to shape any noise that already exists in a signal. SSR achieves the linearity by using preexisting noise in the signal or by explicitly adding noise to the signal before the 1-bit quantization.

SDM is a well-known subject matter and will therefore not be discussed any further in this thesis.

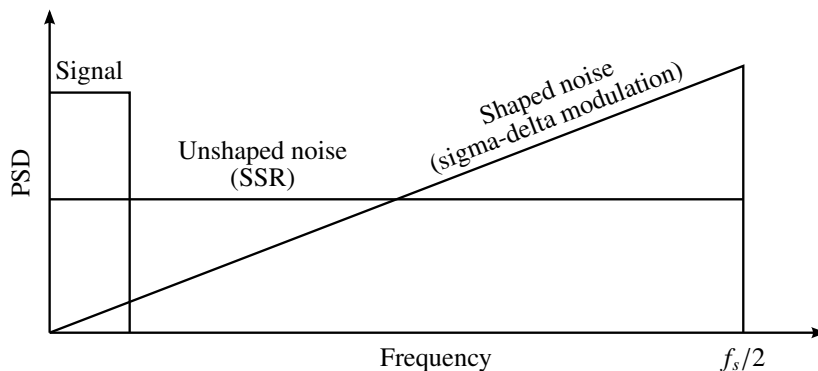


Figure 4.12: Rough sketch comparing shaped and unshaped noise.

4.5 Threshold Levels

The threshold level in an SSR system is usually set to the DC value of the input signal.

A typical parallelized SSR system has all threshold levels set to the same value. For SNR values lower than about 0 dB, this is in fact optimal, but for higher SNR values, multiple threshold levels distributed more like in an ADC is the most optimal [McDo 06a §] [McDo 06b §, ch. 8].

If the SNR is too high, so that SSR does not operate properly, we can set the threshold to a value near the signal value and then do SSR-mode sampling there. The threshold value can be found by using e.g. a quick threshold sweep, binary search, or signal tracking [Papa 98 §]. This technique can only focus on one signal value at a time, so it does not work well for capturing many successive samples, such as in a radar system.

4.5.1 Swept-Threshold Sampling

In the master's thesis of the author [Hjor 06 §], a technique we named “swept-threshold sampling” [Hjor 07 §] was introduced. This is basically SSR with the addition of sweeping of the threshold level, which is much like parallelized SSR with different threshold levels. This makes the system more ADC-like.

To separate between the different modes, we use the terms Stochastic Resonance (SR) to describe a system with a fixed threshold and Swept Threshold (ST) otherwise. Since having one threshold level is optimal when there is much noise and having multiple levels is optimal when there is little noise [McDo 06b §, ch. 8], SR is suitable for low SNR, and ST is suitable for high SNR. Note that ST has an inefficiency which SR does not have. The reason for this is that when the threshold is outside the noise range, no signal integration occurs.

It is interesting to note the difference between SR with uniform noise and ST. The noise values and the threshold values have the same effect in the systems, so since they are uniformly distributed in both types of system, one might think that the systems should behave the same. The critical difference, though, is that the sweep is deterministic, whereas the noise is random. Another way to put it, is that the noise might be white, whereas the sweep has a different type of spectrum.

In an ST system, noise is usually not beneficial, as it is in SR systems. If there is no noise, we will know exactly which two threshold levels the signal lies between. If there is noise, though, the readout becomes more uncertain. In the latter case, the system essentially utilizes a combination of ST and SR. The noise will reduce the demands on how small the step size needs to be, though, since the RMS amplitude of Gaussian dithering noise generally does not need to be larger than the quantization steps [Widr 96 §]. Since the sweep might be performed by a DAC, we can therefore use a simpler one with fewer bits of resolution if there is enough noise in the input signal. The requirements for the 1-bit quantizer might also get relaxed by noise in the input signal, since it then does not need to be able to precisely distinguish values that are very close to each other.

ST systems do not need to know much about the noise, whereas SR systems require detailed knowledge about the noise in order to reconstruct the signal properly. This means that ST is robust against variations in the noise characteristics.

There is some earlier work that discusses similar techniques, both SR, ST, and a sweep-then-SR combination for WSN applications [Papa 98 ⑤]. There is also some later work that has some similarities with ST [Barb 10 ⑤].

4.6 Integration by Oversampling or by Parallelization

Integration, i.e. averaging of multiple readouts of the same input signal, can help improve the output SNR of a system. Two ways of doing integration for an SSR system is oversampling and parallelization, as shown in the block diagrams in figure 4.4 and 4.5. In this section, we will discuss the properties of the two methods.

An oversampled SSR system will have the properties described in the following. Firstly, it will get some general benefits from oversampling, as described in section 4.7. Secondly, colored input noise can be used to improve the system performance, as described in section 4.13. Finally, if the system adds distortion error, some or all of it might be located outside the signal band, thus making it easy to filter out, as described in section 4.9.7.

A parallelized SSR system, on the other hand, will have a different set of properties. Firstly, such a system is easier to analyze than an oversampled system since the samples can be considered independent of each other and we therefore do not need to take into account the time dimension. This is something we exploit for the analysis in section 4.10. Secondly, just a summation of 1-bit values, not any low-pass filter, is needed after the quantization, which could make such a system easier to implement. Finally, the system only needs to run at the Nyquist rate of the input signal instead of at a possibly much higher rate.

Note that oversampling or parallelization just creates more available samples and that the actual integration is done by the low-pass filter or the summation, respectively.

The noise power is usually reduced by n , the integration factor, and the noise amplitude is therefore usually reduced by \sqrt{n} . This means that integrating four times more gives half the noise amplitude, or in other words 6 dB or 1 bit better performance. In an SSR system with oversampling, doubling the sampling rate while keeping the noise amplitude per sample unchanged, will spread the noise power over twice the bandwidth. This means that the noise PSD will be halved and thus the SNR improved by 3 dB. In an SSR system with parallelization, doubling the number of elements will halve the noise power in their sum, thus also here improving the SNR by 3 dB.

There might be a limit to how much we can integrate in an SSR system, though, since the expectation value at the output of a 1-bit quantizer might not be a completely linear function of the input signal. In this case, increasing the noise will actually improve the linearity so that we can integrate much more. This is discussed further in other parts of the current chapter.

Note that in dithered images such as those in figure 3.2, the eyes will integrate spatially so that we essentially are able to experience a low-pass filtered version of the image, even though we can see each pixel clearly.

4.7 General Benefits of Oversampling

PDM, such as SSR and SDM, requires a much higher sampling rate than the Nyquist rate in order to achieve high SNR, or in other words, sufficient oversampling is required in these systems. Regardless of its role in PDM, oversampling also has some beneficial properties by itself, which will be described in this section. Since PDM systems already oversample the signal, these benefits can be reaped without additional cost and are thus a bonus feature of PDM systems.

Sampling preserves all information about a band-limited signal, but it creates some practical problems with conversion and processing. Oversampling mitigates these problems. When working with a sampled signal, it is assumed that between the samples, the signal is reconstructed using the ideal sinc filter. This must be taken into account when performing signal processing. In the frequency domain, the reconstruction filter is a brick-wall filter, and in the time domain, it extends far both forwards and backwards in time. A consequence of this being such a demanding filter, is that implementing it can be expensive. The reconstruction requirement is therefore often ignored when doing signal processing on sampled signals. This is not a problem in many cases, though, but sometimes it will cause performance degradation. These problems are discussed further in the paragraphs below. By oversampling the signal, i.e. by sampling at a higher frequency than required by the Nyquist-Shannon sampling theorem, the behavior of the signal becomes more continuous-time-like, such as that the reconstruction will be “built in”. The challenges associated with sampled signals will therefore become less pronounced.

The phase and frequency response of Nyquist-rate filters are significantly different from analog filters in the top quarter of the frequency range. Oversampling mitigates this problem [Case 93 ☺].

Nonlinear processing such as audio distortion effects, which create high-frequency harmonics, also benefit from oversampling since it prevents harmonics up to $f_s/2$ from folding down into the band of interest [Case 93 ☺].

Another advantage of oversampling is that sharp anti-aliasing and reconstruction filters can be moved from the analog to the digital domain [Stew 98 ☺]. In a Nyquist-rate system, the anti-aliasing filter must filter out all frequencies above $f_s/2$, while still letting through frequencies as high up to this same point as possible. In an oversampled system, f_s is much higher, and the analog anti-aliasing filter still has to remove frequencies above $f_s/2$, but in this case it only needs to let through frequencies up to the original $f_s/2$. This allows its roll-off to be much slower. A digital anti-aliasing filter can then be used to perform the sharp cut-off required to remove the remaining out-of-band noise. The reconstruction filter in a DAC system can be handled in a similar way.

An “analog \rightarrow Nyquist-rate sampled \rightarrow analog” system will have some latency because of the sharp filters required. By using a higher sampling rate, the latency of the system can be reduced.

When we oversample, the in-band quantization noise will be reduced by being spread over a wider band. This works as long as the quantization-noise amplitude stays constant, which it does for multi-bit ADCs and SSR systems where the input-noise amplitude is independent of the sampling rate.

When we oversample and the noise PSD is constant, we might sample a lot of noise, and the

full-band SNR, or in other words the SNR per sample, will become low. This lowers the quality requirement for the sample values, which means that we can perform coarse quantization and that additional noise is not a big problem. This is what enables the use of the coarsely quantized SSR systems.

4.7.1 Oversampling in the Retina

Oversampling can also be seen in biology, where the retina performs filtering on an oversampled signal instead of on the analog optical signal. In the fovea, the sharp central vision, of the primate retina, optical blur ensures proper anti-aliasing filtering before the image is sampled by the photoreceptors [DeVr 97 ⑤, p. 2058]. These photoreceptors are connected one-to-one to ganglion cells, each of which has a long axon that extends into the brain. For the peripheral vision, though, there are fewer ganglion cells than photoreceptors, which means that the photoreceptors oversample the signal and that the ganglion cells then decimate it. Since it is the ganglion cells which connect to the brain, it is the density of these cells which determines the visual resolution. The optical blur is about the same in the periphery as in the fovea, so a simple downsampling would result in aliasing. However, the ganglion cells do spatial low-pass filtering, i.e. anti-aliasing filtering, which prevents this.

The ganglion cell receptive fields, i.e. sensitivity profiles, in the rabbit retina have been found to be very close to two-dimensional Gaussian functions with a slight ellipticity [DeVr 97 ⑤, fig. 4]. The receptive fields are arranged in an approximate hexagonal pattern, with a 2-sigma center-center spacing, or in other words with their 1-sigma borders touching. This arrangement causes the spatial frequency response of the receptive fields to be close to zero above the Nyquist frequency of the hexagonal pattern, thus ensuring good anti-aliasing filtering before the downsampling [DeVr 97 ⑤, fig. 10]. So to sum up, the processing chain goes like this: Optical anti-aliasing filtering → sampling (oversampled) → neural anti-aliasing filtering → downsampling (part of the neural filtering).

So as we can see, the retina is a system that employs oversampling. But why does it do that? One way to do the anti-aliasing filtering would be to have different degrees of optical blur for the fovea and the peripheral vision, but this might perhaps be difficult to achieve and might therefore not be a viable solution. Another possible solution would be to have big non-overlapping photoreceptors with uniform sensitivity, but that would not give very good anti-aliasing filtering since the spatial frequency response of the photoreceptors would then be sinc-like due to the brick-wall-like spatial response. However, with the use of oversampling, overlapping two-dimensional Gaussian sensitivity profiles can be realized. As explained above, these provide good anti-aliasing filtering. Another benefit of the oversampling is that different ganglion cell classes can have different spatial resolutions and anti-aliasing filters [DeVr 97 ⑤, fig. 6]. With a fixed anti-aliasing filter, changing just the resolution would either cause aliasing or an unnecessarily blurry image. Thanks to the oversampling, though, an appropriate filter can be used for each of the resolutions. The different ganglion cell classes in the retina perform different image processing operations, and in the rabbit retina, 11 such classes have been identified.

Oversampling is also being used in some CMOS cameras [Vuor 13 ⑤], which means that the image is acquired at very high resolution and then scaled down to the target resolution.

This should make it possible to use digital anti-aliasing filters that are better than the optical anti-aliasing filters in conventional image sensors. It is unclear whether this opportunity is actually being exploited to its fullest or if the oversampling is primarily being used for other properties such as zooming, though.

Note that this oversampling of images followed by anti-aliasing filtering is similar to SDM ADCs, which also oversample and then perform anti-aliasing filtering on the oversampled signal, as described in the previous section. The motivation for using this technique might be slightly different in the two cases, though.

4.8 Decoding

The DC transfer function of an SSR system is the Cumulative Distribution Function (CDF) curve of the input noise, as discussed in section 4.11. When the input noise is Gaussian, the DC transfer function will therefore be approximately linear for a small input signal range and nonlinear for larger signals, as seen in figure 4.18. For uniform noise, the CDF will be a straight line, though, so the DC transfer function will then be completely linear, except that it is clipped at the extremes, of course, as seen in figure 4.19.

Let us assume that the SSR system uses parallelization, not oversampling. We will not be looking into oversampled SSR systems here since that is probably more complicated. The output from the parallel 1-bit quantizers are summed, which gives us a discrete number. This number might then be converted to an analog value again. This is known as decoding. In order to compensate for the nonlinearity of the DC transfer function, we will have to use a better decoding scheme than a simple linear one. There are many ways to choose an optimal decoding [McDo 06b §, ch. 6 and 7], but we will not go any further into this here, since it has already been investigated in literature. An important thing to remember, though, is that an optimal decoding might be much more complicated to implement than a linear one, so this kind of compensation is a tradeoff with system complexity.

4.9 SNR Metrics

The SNR of a signal is measured in several different ways in this thesis. It can therefore be useful to discuss the different methods to clarify the differences. The problem at hand is how to measure the SNR of a signal which has noise and distortion, and in particular, how to separate the signal from the noise and distortion.

An alternate problem definition is how to set the output gain of a system so that the signal part of the output has the same amplitude as the input signal. Again, this requires a well-defined way of identifying the signal part. Such a problem definition is useful for the systems in this thesis which are based on 1-bit quantizers since they in their primitive form simply output the values 1 and -1 .

Note that Signal-to-Noise-And-Distortion ratio (SINAD) might be a more correct term than Signal-to-Noise Ratio (SNR) here since we include distortion together with the random noise. There are, however, several different definitions of SINAD, so the term is a bit ambiguous. Furthermore, the term is less used than SNR, and an expression we often use in this thesis, “SNR loss”, is practically never written as “SINAD loss”. We will therefore use the terms “SNR” and “SNR loss” throughout this thesis.

The SNR metrics discussed here are for analytical or numerical evaluation of systems, not for practical SNR measurements, since the noiseless input signal is used in the calculation of the SNR.

4.9.1 Measurement Setup

The SNR measurement setup we use here is shown in the block diagram in figure 4.13. A noiseless input signal is sent through a system under test which distorts the signal and adds random noise. Instead of looking at the distortion as a nonlinear transform of the original signal, we can look at the distortion error, i.e. the deviation from the correct output signal, as a type of additive output noise. Together with the random noise, this makes up the total output noise. We define the output signal, i.e. the signal part of the SNR, to be equal to a scaled version of the input signal. The total output noise, or output error, is the total output minus the output signal. The block diagram also shows two controllable gains. We will get back to these later. So to sum up, total output = output signal + output noise = scaled input signal + distortion error + random noise, or as shown in the block diagram, output noise = total output - output signal = distortion error + random noise, and output signal = scaled input signal.

The SNR is the ratio of the output signal power to the output noise power. The output noise power is here the same as the Mean Squared Error (MSE) at the output.

We primarily consider a sampled system where the signal uses the entire bandwidth, so that we do not get an in-band and an out-of-band region in the spectrum, which could complicate the analysis. By also assuming that the random noise is white, each sample will be independent, and we therefore do not need to consider the time dimension. All we have to do is to analyze the probability distributions of scalar values. We will later briefly discuss oversampled systems, though, where the distortion harmonics can get pushed out of the signal band.

The three similar plots in figure 4.13 illustrate what happens to an input signal that gets

passed through an SSR system which has input noise $\sigma_N = 0.5$, $n = 20$ parallel thresholders, and an output value range of $\pm \sqrt{\pi/2} \cdot \sigma_N$. For a given input signal value, the output has a probability distribution which is here a Probability Mass Function (PMF) making up one vertical segment of the gray shape in e.g. the upper left plot. The reason it is a PMF and not a PDF, is that the SSR system will produce discrete output values. The expected value of the distribution, the black curve, is the distorted signal, i.e. the output signal plus the distortion error. The distortion error is therefore the difference between the output signal, i.e. the scaled input signal, shown as a red curve, and the distorted signal. The random variation around the expected value is the random noise, and the variance of the distribution is therefore the power of the random noise. The total power, or MSE, of the distortion error and the random noise, i.e. the noise part of the SNR, is simply the sum of the power of each of these two components. The reason a straight-forward power addition can be used here, is that the two components are uncorrelated, as discussed on page 100 in section 4.11.

4.9.2 Gains

The only thing that remains now is to choose the gain which is used when defining the output signal as a scaled version of the input signal. This task requires careful consideration since it can be done in several different ways which may look quite similar but which could yield very different results in some cases.

Said controllable signal-part gain can be seen in the block diagram in figure 4.13, and an additional controllable output gain can also be seen. If both of these gains are changed by the same factor, the SNR will of course be unchanged, so it might seem like there is an unnecessary degree of freedom here. It turns out, however, that when trying to find the correct gain setting by sweeping the gain, the end result depends on which of the two gains are being swept. This will be explained in detail below. Note that we will sometimes treat the output gain as a part of the system under test rather than just as a part of the setup for measuring the SNR.

An important question is when to set the gains. This can be done when we make an SNR measurement, in real time by the system under test, or just once in advance. When the input signal characteristics, especially the amplitude, changes, the amount of distortion can change so that the correct values for the gains will change too. This means the gain settings should in principle be updated continuously, but that is not always the correct thing to do.

If we are not interested in the absolute value of the output signal, i.e. if the amplitude of the output signal does not matter, we can use an SNR metric which always uses up-to-date gain settings. Such an SNR metric will find out what the output signal amplitude is for each new output sequence instead of e.g. just assuming that the output signal is exactly identical to the input signal regardless of varying compression by the system. This dynamic adjustment of the gains is correct to use when we only care about the shape of the output signal and not its amplitude.

If we are interested in the absolute value of the output signal, though, the system itself must produce the correct absolute value, so in this case we can not change the gains just for the SNR measurement since that would not correctly reflect the performance of the system. We will therefore have to use a naive SNR metric where the output noise simply is the deviation

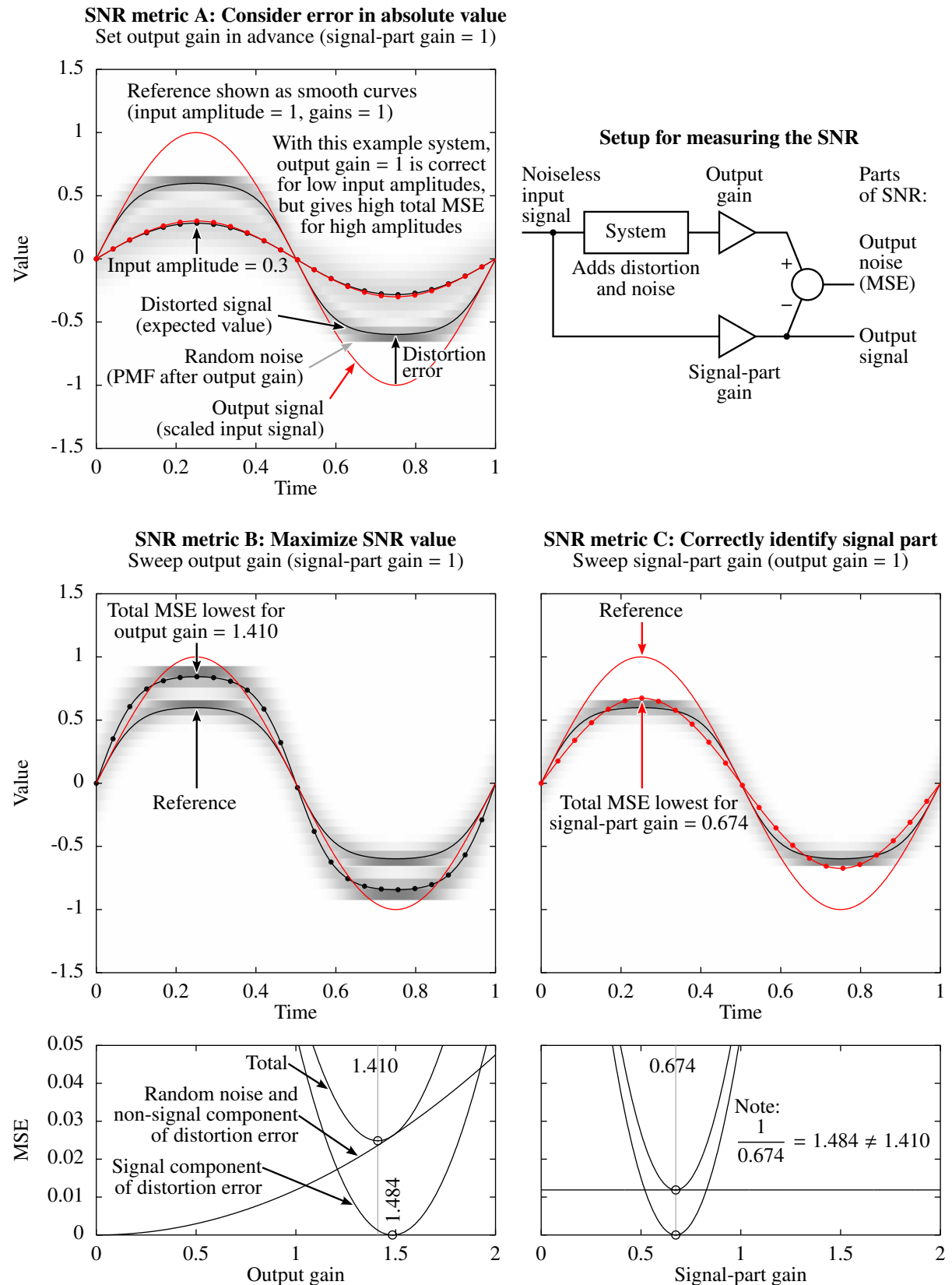


Figure 4.13: Illustrations of different SNR metrics.

from the correct absolute value, i.e. where the signal-part gain is set to 1 and the output gain is controlled by the system under test. This naive SNR metric will register an incorrect output signal amplitude as a large error, i.e. strong noise, like it should, whereas other SNR metrics might not consider this to be an error at all. If the system is not able to continuously tune the output gain as the input characteristics change, we will just have to set the output gain in advance and then live with a sometimes incorrect gain value.

We use three different methods for finding the gain values, which we have called SNR metrics A, B, and C. In the process of finding the values, we may sweep one gain setting or the other, but note that after the gain values have been found, it is possible to change both gains by the same factor in order to e.g. always keep the signal-part gain at 1 so that the output signal gets the same amplitude as the input signal.

4.9.3 SNR Metric A: Consider Error in Absolute Value

Instead of updating the gains for each SNR measurement, we simply set the output gain in advance in this naive SNR metric. The signal-part gain is kept at 1. This SNR metric therefore defines the output signal to be equal to the input signal and thus the output noise to be the deviation from the absolute value, i.e. the unscaled value, of the input signal.

Forcing one of the gains to 1 is not a requirement, though, as any two gain values could actually be used, as long as they are set before the system under test is used and the SNR measurement is made.

The output gain value could e.g. come from a small-signal analysis of the system or from SNR metrics B or C performed ahead of time with an assumed input signal distribution. As long as conditions match those used to calculate the gain with another SNR metric, SNR metric A will be identical to that metric.

One reason for choosing this SNR metric is that it is very simple. Another reason is that it is the correct way to measure the SNR when we are interested in the absolute value of the output signal and we define the noise to be the deviation from the correct absolute value. In this case the gains can not be updated by the SNR metric since the system under test will have to do that itself in order to produce the correct absolute output values. So here the output gain is a part of the system under test, either as a fixed value or as a parameter being updated in real time by the system as the input characteristics change.

The upper plot in figure 4.13 shows a setup with both gains set to fixed values of 1. The system under test itself has an output value range of $\pm \sqrt{\pi/2} \cdot \sigma_N = \pm \sqrt{\pi/2} \cdot 0.5 \approx \pm 0.627$, as described earlier. That range, i.e. the native output gain of the system, is set according to a small-signal analysis of the system. So if we see this native output gain, which is part of the system under test, and the external output gain, which is part of the SNR measurement setup, as one combined value, we can say that the output gain is controlled by the system under test and set to a fixed value according to a small-signal analysis of the system. As we can see, the system performs well for an input amplitude of 0.3, but as we can also see, it will have a large MSE for an input amplitude of 1. The MSE will be so large not just because of the distortion itself but also because the output gain does not get updated as the strength of the distortion changes. In other words, this SNR metric registers that the system under test produces quite incorrect

absolute values in the latter case, which makes it a good SNR metric when the absolute value is of interest.

An example application of this SNR metric is for a typical SSR system as shown in figure 4.6. The quantizer output levels are set to $\pm \sqrt{\pi/2} \cdot \sigma_N$ in such a system, which makes the expected value at the output equal to the input signal, at least for small signals. SNR metric A can therefore be used directly on this system, where it will use the deviation from the correct absolute value as the noise part of the SNR. When the input signal amplitude is low, there will be little or no distortion error, just random noise, so this SNR metric will then be identical to SNR metric C, which is described below. This is therefore a metric for consideration of the absolute value which also, under certain conditions, can be used as a simplified implementation of SNR metric C. Because SNR metric A is simple, considers the absolute value, and can sometimes be used as a substitute for SNR metric C, we use it in the simulation of 1-bit quantizers in section 4.14.

4.9.4 SNR Metric B: Maximize SNR Value

In this SNR metric, we sweep the output gain in order to find the value that gives the lowest MSE. The signal-part gain is kept at 1. This maximizes the reported SNR value.

The reference setup, shown as smooth curves in the three upper plots in figure 4.13, is the same in all three plots and has both gains set to 1. SNR metric B sweeps the output gain and SNR metric C sweeps the signal-part gain in order to bring the scaled input signal and the total output closer together, as can be seen in the two middle plots. This reduces the output error, i.e. the MSE or the output noise, and essentially identifies the signal part of the output. Somewhat surprisingly, though, the two methods give different end results.

The total MSE is plotted against the output gain in the lower left plot in figure 4.13. It is useful to decompose this total MSE, or output error, into constituent parts. The output error consists of random noise and distortion error, and the distortion error can be further decomposed into a scaled version of the input signal, i.e. the signal component, and the remainder, i.e. the non-signal component. The signal component can be identified by calculating the correlation between the input signal and the distortion error. As the output gain changes, the amplitudes of the random noise and the non-signal component scale accordingly, so we plot these similarly-behaving components as one curve, seen as the right half of a parabola going diagonally across the plot. The signal component of the distortion error behaves differently, though, since it is zero when the signal component of the distorted output signal is equal to the scaled input signal, and non-zero otherwise. If the scaled input signal subtracts the wrong amount from the distorted output signal, there will be a signal component left in the distortion error, either with a positive or a negative sign, which will make the MSE increases if there is any deviation from the correct output gain value in either direction. This curve can therefore be seen in the plot as a parabola which has a minimum of zero which is located at a non-zero value of the output gain. The total MSE is a straight-forward sum of the two MSE components since the power of two uncorrelated signals can simply be added together in order to find the total power. Note that an SNR plot could not have been decomposed like this since the output errors would then have been in the denominator. This is the reason we plot the MSE instead of the SNR.

When the output gain is 1.484, there is no remainder of the signal in the output noise, so the

signal component of the distortion error has an MSE of zero, as can be seen in the plot. This is, by the way, the condition we are looking for with SNR metric C. The total MSE, though, has its minimum at a lower output gain value of 1.410, which can also be seen in the plot. In other words, the reduction in output gain from 1.484 to 1.410 gives us a signal component in the output noise, which will contribute to the total noise power, but the amplitude reduction for the rest of the noise makes the total noise power lower anyway. This is why minimizing the total MSE by sweeping the output gain is not the same as identifying the actual signal part of the output.

This SNR metric maximizes the reported SNR value. That is because the signal power is constant due to the constant signal-part gain of 1, so that $\text{SNR} = (S / \text{total MSE})$ is maximized when the total MSE is minimized. The metric does not correctly identify the signal part of the output, though. In other words, because of the way the output noise is defined, it can contain some of the signal. This noise component will be a scaled and inverted version of the input signal. The maximized SNR value can therefore be said to be a somewhat false SNR metric, at least for some applications, so it must be used with care. An example of its special behavior is that the maximized SNR value will never be below 0 dB since an output gain of zero would give exactly that value.

This SNR metric is well suited for applications where we are interested in the absolute value of the output signal and we want to minimize the deviation from the correct absolute value. If we set the output gain of the system according to this SNR metric, the output signal may not get the correct amplitude, but the MSE at the output will be minimized, which might be exactly what we want.

When we use this SNR metric to assess the absolute value of the output from a system, it is important to be aware that the metric is intended for theoretical evaluation of the system in advance of actual system usage. The evaluation is done for a given input signal distribution, and the result is an output gain value and an SNR value. The system itself can then be configured with this output gain value to give it the desired properties. When we later want to measure the SNR of a finite-length output sequence, we can simply use SNR metric A. This gives a kind of instantaneous SNR which will vary from time to time, unlike the theoretical SNR. If SNR metric B was used directly on a finite-length sequence, though, the output gain would get tuned to the optimal value for that exact sequence. Such an SNR measurement would not reflect the actual performance of the system since it is the system itself which should produce the correct absolute values. We therefore use SNR metric A, which does not change the gains, but just observes the absolute-value performance of the system.

An example application of this SNR metric is in finding the optimal quantizer output levels for a 1-bit quantizer which has a white Gaussian input signal and no input noise. SNR metric B gives quantizer output levels of $\pm \sqrt{2/\pi} \cdot \sigma_{\text{signal}}$ in this case. This type of system, in particular optimized with SNR metric B, is discussed in literature [Max 60 ⑤] and also shown in figure 4.46. The choice of SNR metric B optimizes the system for the criteria of this metric, so if other system properties are desired, other SNR metrics should of course be used.

4.9.5 SNR Metric C: Correctly Identify Signal Part

In this SNR metric, we sweep the signal-part gain in order to find the value that gives the lowest MSE. The output gain is kept at 1. This identifies the actual signal part of the output, unlike the more artificial definition of the signal part used in SNR metric B.

Another way to look at it is that we try to subtract as much of the signal as possible from the output, i.e. minimizing the remaining output noise after we subtract the signal part from the output. Instead using these sweep-based methods, we could calculate the correlation between the input signal and the output in order to identify the signal part of the output.

See the section above about SNR metric B for an explanation of the plots in figure 4.13, which also illustrate SNR metric C.

The signal component of the distortion error is the only output noise component which changes as the signal-part gain is swept, so the total MSE has its minimum at the same signal-part gain value as the minimum, i.e. the zero value, of the MSE of the signal component of the distortion error. This can be seen in the lower right plot in the figure. When this zero value is reached, the scaled input signal will be exactly equal to the signal component of the distorted signal since that leaves no signal component in the distortion error. This condition is shown by the dotted curve in the middle right plot. So when we minimize the total MSE, we find the actual amplitude of the signal part of the output. Note that also here, as for SNR metric B, the expression is $\text{SNR} = (S / \text{total MSE})$, but unlike in SNR metric B, the signal-part gain and the signal power are not constant during this sweep. Therefore, the SNR will not be maximized when the total MSE is minimized, as it is in SNR metric B. This shows why SNR metrics B and C, i.e. sweeping the output gain and sweeping the signal-part gain, give different end results.

This SNR metric finds the actual amplitude of the signal part of the output. When there is no distortion, the signal part of the output, as defined by this metric, is exactly equal to the expected value of the output signal. The metric does not find signals which are not really there, like SNR metrics A and B do, so if, for example, there is no trace of the input signal in the output, this metric will report an SNR of $-\infty$ dB.

This SNR metric is suitable for when we are not interested in the absolute value of the output, i.e. when the scaling of the output signal is not important, but when we want to work with the actual signal part of the output instead of an artificially defined one. If we are interested in the absolute value and it is more important that we get the true output signal than it is to have a minimized MSE at the output, we could use SNR metric C anyway, though. We would then use SNR metric C to set up the system and then subsequently use SNR metric A.

In the section above, we stated that SNR metric B should be used for theoretical analysis of a system and that SNR metric A should then be used for instantaneous SNR measurements on finite-length output sequences. SNR metric C can be used like this too, as just mentioned, but since we are usually not interested in the absolute value of the signal with this metric, we can let the SNR metric tune the gains for each instantaneous SNR measurement. Note that such instantaneous measurements with SNR metric C will have different amounts of noise each time and that some of the noise could be detected as the signal part. As an example, a system which ignores the input and simply outputs white Gaussian noise would seem to have some signal in the output most of the time, so that the measured SNR would often be a low finite value instead

of $-\infty$ dB.

An example application of this SNR metric is for measuring the SNR of an SSR system as a function of input noise, as shown in figure 4.21.

4.9.6 Examples of the Differences

Figure 4.23 shows the three different SNR values plotted against input noise for a 1-bit quantizer, which clearly demonstrates how the three SNR metrics can yield very different SNR values.

The three SNR metrics each give different results when used to configure the output gain of a system, too. Figure 4.24 shows what the quantizer output levels of a 1-bit quantizer should be according to the three different SNR metrics in order to give the output signal the same amplitude as the input signal. As we can see in the figure, the three SNR metrics give very different results. In the region of negligible input noise, for instance, SNR metric C gives quantizer output levels of $\pm \sqrt{\pi/2} \cdot \sigma_{\text{signal}} \approx \pm 1.25 \cdot \sigma_{\text{signal}}$ and SNR metric B gives quantizer output levels of $\pm \sqrt{2/\pi} \cdot \sigma_{\text{signal}} \approx \pm 0.80 \cdot \sigma_{\text{signal}}$. So in other words, the configuration of the system will be quite different depending on which SNR metric is used when optimizing it.

4.9.7 Oversampling

So far we have considered sampled systems where the input signal uses all of the available bandwidth, i.e. systems that are not oversampled. By making the system oversampled, though, it gets some new and interesting properties. First of all, we should then measure the SNR in the signal band instead of for the full bandwidth. If the random noise is white, it will be spread across the full bandwidth so that only a small part of it falls inside the signal band. This may not represent any kind of real SNR improvement, though. Something that could, however, give an actual increase in the SNR, is when some or all of the distortion error power is pushed out of the signal band. Such out-of-band distortion harmonics can be seen in figure 4.52.

When the input signal is narrowband, all of the distortion harmonics may be located outside the signal band. Note that higher harmonics could alias back into the signal band, though, and that the distortion error might have a fundamental frequency component too. When the input signal has a low-pass spectrum characteristic, however, there will probably always be overlap between the signal and the distortion harmonics.

The SNR metrics described above may not be directly transferable to the case of oversampled systems. For example, SNR metric B, which maximizes the SNR value, should now probably consider the in-band SNR, so its implementation might therefore need an update.

Be aware of the difference between the theoretical long-run distribution of an input signal and the instantaneous distribution, so to say, of a finite-length sequence of it. As an example, white Gaussian noise filtered to a single frequency in a simulation will have a Gaussian distribution in the long run, but during a single simulation iteration, it will just be a sinusoid with a random amplitude and phase. When a system is e.g. designed for a Gaussian input signal but gets a sinusoidal input signal during a given simulation iteration, its performance might suffer, which could e.g. give in-band distortion error even for narrowband signals. In other words, the distortion will then not be limited to just harmonic components.

Section 4.14 shows the spectra from many system simulations, most of which use oversampling.

4.9.8 Distortion Compensation

Instead of using a scaled version of the input signal as the output signal, we could just use the distorted output signal and assume that it will be corrected at a later point in the system. Alternatively, we could add a distortion compensator after the system under test. This might complicate the SNR metrics and would require compensation circuits in real systems. When there are no such circuits, though, we should probably just stick with the type of SNR metrics we have discussed so far.

4.9.9 Limitations of the SNR Metric

Note that the SNR metric in general and the presented SNR metrics A and B in particular can sometimes behave unexpectedly and that other types of metrics might yield other results. As an example, information theoretic metrics have been used in literature to analyze SSR systems [McDo 06b ⑤].

One example of the unexpected behavior is that SNR metrics A and B somewhat artificially define what the output signal is, which means that a signal could be detected where there is none at all. A second example is that SNR metric B never yields a result below 0 dB. An example of where the SNR metric can be particularly deceptive, though, is when it reports a higher SNR at the output of a system than at its input. This can be seen in figure 4.51.

A situation where the SNR of the input is improved by a system, is if we have a sinusoidal signal with white Gaussian noise and we send this input through a system which clips it at the top and the bottom at the amplitude level of the sinusoid. If we are interested in the absolute value of the signal, everything outside the amplitude range is guaranteed noise, so the clipping will reduce the MSE at the output and the system will thus increase the SNR. Another situation where the SNR can be improved, is if we have a Binary Phase-Shift Keying (BPSK) transmission, i.e. a sequence of +1 and -1 values in the baseband, with noise on top of it. Thresholding that sequence could improve the SNR, but the BER would not improve. This shows how an improved SNR is not necessarily quite as impressive as it sounds, and it also demonstrates how SNR is different from measures of information content. The thresholding and the SNR improvement could be useful for maintaining signal quality in a chain of relay stations, though, so this kind of SNR improvement can still be useful in some cases. It is up for debate how real or useful the SNR improvements are, but the bottom line is that SNR values must be interpreted with care.

4.10 Small-Signal Analysis

A system using Suprathreshold Stochastic Resonance (SSR) to capture an input signal distorts it in the process and thus loses information about it. We will now analyze how an SSR system compares to an Analog Average (AA) system in terms of the standard deviation of the recovered signal when the input signal is constant. The analysis will be done for the asymptotic case where the integration factor n is very high and the input SNR is very low. Note that we want the SNR to be low in order to linearize the system and that a large n will greatly reduce the output noise and might thus expose nonlinearities. Therefore, n can not be too high for a given SNR.

The system is shown in figure 4.14. x is the input signal, and we assume that its value is always close to 0 so that the behavior of the system does not change with x . If the system behavior had changed with x , we would have needed to know the PDF of x in order to weigh the changing effects of the system according to how often they are encountered. Since x is small and thus the system behavior constant, we do not need to know anything further about the PDF of x , though. We can just find the system behavior for $x = 0$, and then we also know its behavior for all small values of x and all low-amplitude input PDFs. Further, we assume that the noise N is white and that it has distribution PDF_N with median 0. Then, when $x = 0$, the thresholder will have a probability of $p = 0.5$ of outputting a “1”, and when x changes, p will change too, as illustrated in figure 4.15 and as given by this formula:

$$p = 0.5 + \text{PDF}_N(0) \cdot x \quad (4.1)$$

While x stays fixed at a given value, n output values from the thresholder, each of which is 1 or 0, get integrated in a counter to produce the counter value c , which will have a binomial distribution [W] with:

$$\bar{c} = n \cdot p \quad (4.2)$$

To convert from the count to the value of the recovered output signal x_{SSR} , we use the following formula. Note that this simple formula might not be suitable for large values of x .

$$x_{\text{SSR}} = (c - n/2) \cdot k_{\text{cv}} \quad (4.3)$$

k_{cv} is a count-to-value conversion constant. Since this is a linear equation, it must also apply to the mean values:

$$\overline{x_{\text{SSR}}} = (\bar{c} - n/2) \cdot k_{\text{cv}} \quad (4.4)$$

Expanding:

$$\overline{x_{\text{SSR}}} = n \cdot \text{PDF}_N(0) \cdot x \cdot k_{\text{cv}} \quad (4.5)$$

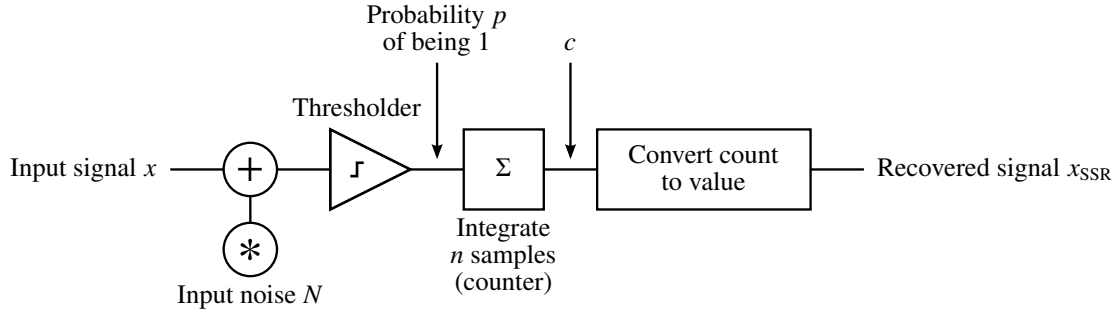


Figure 4.14: An integrating SSR system for sampling a signal.

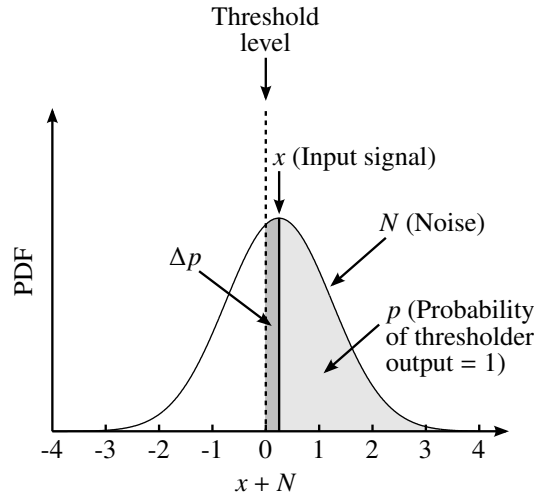


Figure 4.15: Threshold output probabilities.

We want $\overline{x_{SSR}}$, the mean of the recovered signal, to be equal to x , the input signal, so k_{cv} must be:

$$k_{cv} = \frac{1}{n \cdot \text{PDF}_N(0)} \quad (4.6)$$

The standard deviation of the binomially distributed count, σ_c , is given by a known formula, and since $p \approx 0.5$, we get:

$$\sigma_c = \sqrt{n \cdot p \cdot (1 - p)} \approx \frac{\sqrt{n}}{2} \quad (4.7)$$

The standard deviation of the recovered signal, σ_{SSR} , will then be:

$$\sigma_{SSR} = \sigma_c \cdot k_{cv} = \frac{1}{2 \cdot \text{PDF}_N(0) \cdot \sqrt{n}} \quad (4.8)$$

For comparison, the standard deviation of the recovered signal when using analog averaging, σ_{AA} , is:

$$\sigma_{AA} = \frac{\sigma_N}{\sqrt{n}} \quad (4.9)$$

Noise distribution	$\sigma_{\text{SSR}}/\sigma_{\text{AA}}$	$n_{\text{SSR}}/n_{\text{AA}}$	SNR loss of SSR
Trident	$\sqrt{9\pi/302} \approx 0.3060$	$9\pi/302 \approx 0.0936$	-10.29 dB
Gaussian	$\sqrt{\pi/2} \approx 1.2533$	$\pi/2 \approx 1.5708$	1.96 dB
Uniform	$\sqrt{3} \approx 1.7321$	3	4.77 dB
Sinusoidal	$\pi/\sqrt{2} \approx 2.2214$	$\pi^2/2 \approx 4.9348$	6.93 dB

Table 4.1: Performance comparison of SSR and AA.

Where σ_N is the standard deviation of the input noise. The following ratio gives a performance measure of SSR compared to AA in terms of noise in the recovered signal:

$$\sigma_{\text{SSR}}/\sigma_{\text{AA}} = \frac{1}{2 \cdot \text{PDF}_N(0) \cdot \sigma_N} \quad (4.10)$$

So to sum up, we always have $\sigma_c = \sqrt{n}/2$, and the larger $\text{PDF}_N(0)$ is, the larger $\Delta\bar{c}/\Delta x$ is and thus the larger the SNR will be, and, secondly, the larger σ_N is, the worse AA will be and thus the better SSR will be in comparison. So SSR has its best performance when the input noise PDF has a peak at the center but still a large σ_N .

With a white Gaussian input noise having $\sigma_N = 1$, we will have $\text{PDF}_N(0) = 1/\sqrt{2\pi}$, and $\sigma_{\text{SSR}}/\sigma_{\text{AA}} = \sqrt{\pi/2} \approx 1.2533$, i.e. SSR has about 25 % higher noise amplitude at the output than AA in this case. We can compensate for this by integrating more. Since the noise amplitude at the output is proportional to $1/\sqrt{n}$, we have to multiply n with the square of $\sigma_{\text{SSR}}/\sigma_{\text{AA}}$. That gives us $n_{\text{SSR}}/n_{\text{AA}} = (\sigma_{\text{SSR}}/\sigma_{\text{AA}})^2 = \pi/2 \approx 1.5708$, i.e. if we integrate 57 % more with the SSR system than with the AA system, they will both have the same amount of output noise. The output noise ratio $\sigma_{\text{SSR}}/\sigma_{\text{AA}}$, or in other words the penalty of using an SSR system, can also be expressed in decibels, and in this case it is 1.96 dB. This is the quantization loss or SNR loss of the SSR system.

An interesting thing to note is that if we e.g. double σ_N , the $\text{PDF}_N(0)$ value will be halved, and $\sigma_{\text{SSR}}/\sigma_{\text{AA}}$ thus stays constant. Since $\sigma_{\text{SSR}}/\sigma_{\text{AA}}$ only depends on these two variables, the SNR loss will be determined only by the PDF of the input noise, and specifically only by its fundamental shape, not any shrinking or stretching of it.

We will now look at some other input noise distributions. These are shown in figure 4.16. Note that the noise signals shown can not be used directly sample by sample since that gives a predictability to the noise signal, so instead, each noise sample must be picked from a random position on the noise signal curve, or simply be picked according to the PDF. We assume that the noise is white, i.e. that each noise sample is independent of the others. Note that, for the purpose of the $\sigma_{\text{SSR}}/\sigma_{\text{AA}}$ analysis, these different input noise distributions must be preexisting on the input since we have assumed that the SSR system and the AA system receive the same noise. If the noise were explicitly added to help out the SSR system, that would be unfair to the AA system, which just requires as low input noise as possible.

The performance of SSR systems with these different kinds of input noise distributions is shown in table 4.1. We can see that both the uniform and sinusoidal distributions are significantly worse than the Gaussian distribution. We have, however, created a new distribution which we have called “trident”. This distribution was created with the expression for $\sigma_{\text{SSR}}/\sigma_{\text{AA}}$ in

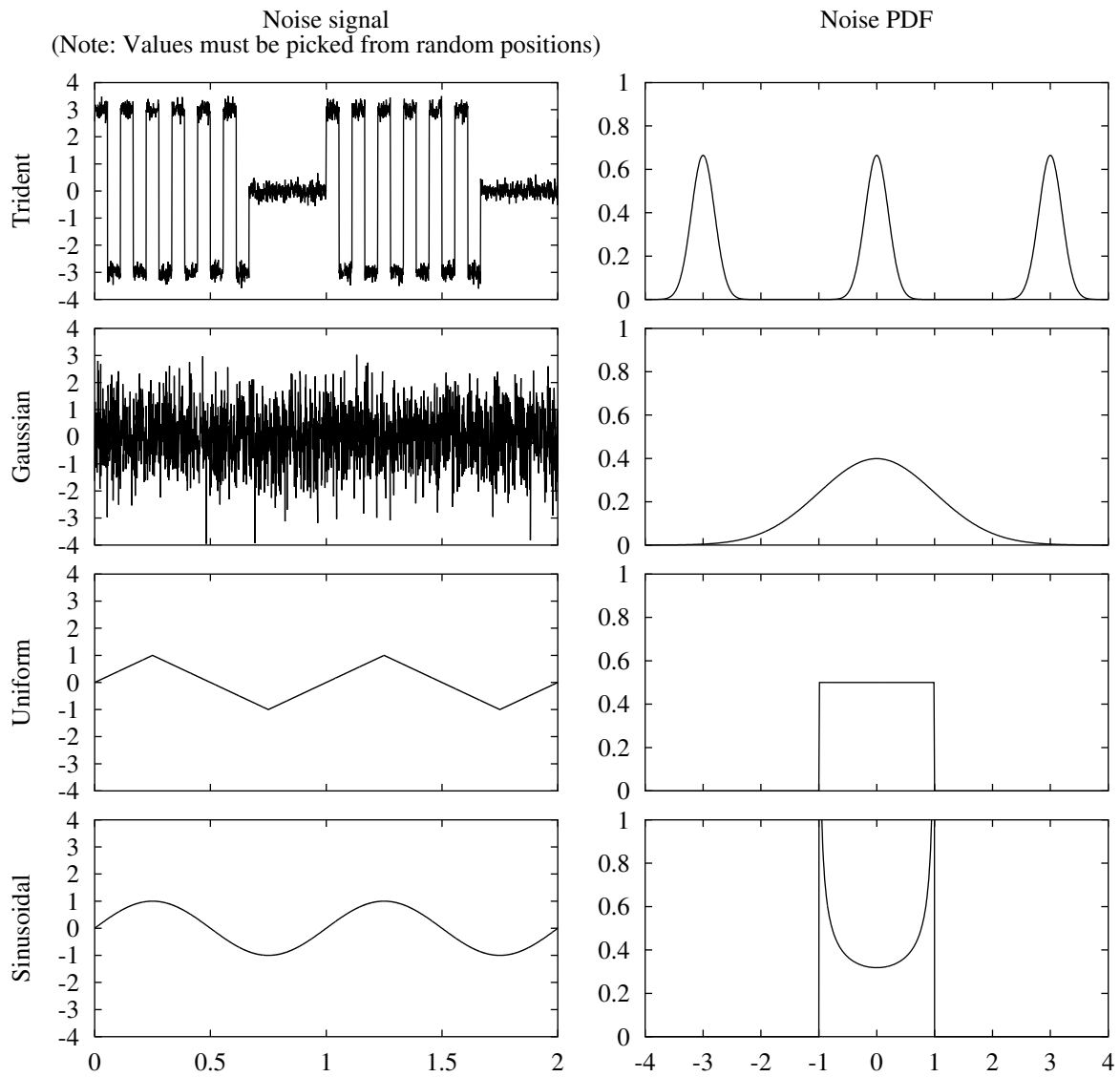


Figure 4.16: Some noise distributions possible at the input of SSR and AA systems.

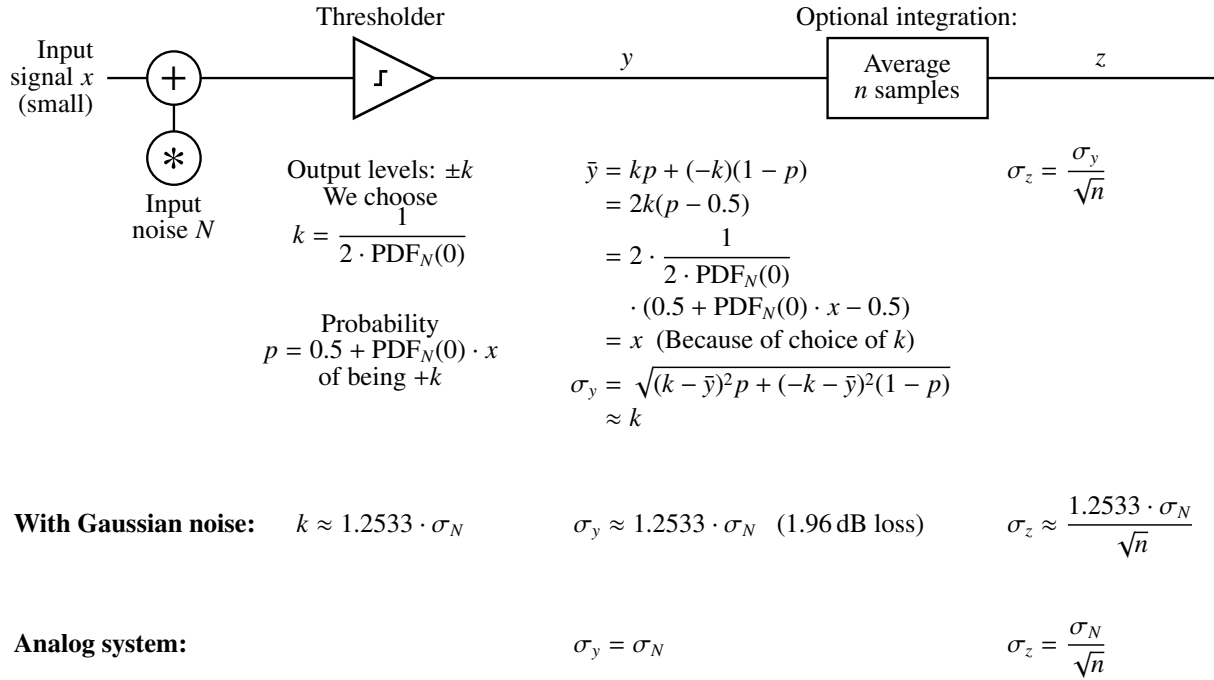


Figure 4.17: Analysis of 1-bit quantization, showing the characteristic 1.96 dB loss.

mind. We wanted the PDF to have a peak at the center while also having a large σ_N , and a PDF consisting of three Gaussians as shown in figure 4.16 fits the bill. In fact, any such bursty noise will do, e.g. intermittent interference. As can be seen from the table, the SSR system performs spectacularly under this input noise condition, with $\sigma_{\text{SSR}}/\sigma_{\text{AA}} \approx 0.3$ and $n_{\text{SSR}}/n_{\text{AA}} \approx 0.1$. This means that the SSR system has less output noise than an AA system or that it only requires a tenth of the integration that an AA system needs in order to achieve the same output noise level. The SNR loss of about -10 dB means that the output SNR is actually improved by switching from an AA system to an SSR system.

The trick with the trident noise distribution is that the extreme values are simply guaranteed 1's or 0's out from the thresholder, which is just like regular lower-amplitude noise to the SSR system. In an AA system these large values are directly averaged into the output signal, so the larger the extremes are, the bigger the noise at the output will be. In an SSR system, though, the strength of the extremes makes no difference.

Again, note that this analysis is for asymptotically large n and low input SNR and that it is the SNR we are looking at at the output. If we use other metrics like e.g. mutual information, some of the results may be different. Furthermore, our basis of comparison is AA, but that might not always be the right choice.

For the case of $n = 1$, i.e. no integration, we show a noise analysis in figure 4.17. This analysis is very similar to the one already shown, so we will not elaborate on it any further. Since we are not integrating here, the system might not exhibit SSR, i.e. the noise might not be beneficial, so “1-bit quantization” is a more correct name for the system than “SSR”. We can, however, add integration as an extra stage at the end of the system as shown in the figure. Then we get the same result as in our first analysis, so this could be an alternative way of analyzing a full SSR system. For the case of Gaussian input noise, we see that the thresholder

output levels must be $\pm k = \pm \sqrt{\pi/2} \cdot \sigma_N \approx \pm 1.2533 \cdot \sigma_N$ in order for the expected value at its output to match the input signal value. So if one has a 1-bit quantizer in a system and it is known that the input has low SNR and Gaussian noise, one can just define the threshold output levels to be $\pm \sqrt{\pi/2} \cdot \sigma_N$ instead of 1 and 0, and then any further processing on the 1-bit signal can be done as if it were simply a low-SNR analog signal, without any need for other kinds of compensation. The threshold output level values will also be the output noise level, i.e. $\sigma_y \approx \sqrt{\pi/2} \cdot \sigma_N \approx 1.2533 \cdot \sigma_N$. The value $\sqrt{\pi/2} \approx 1.2533$, corresponding to 1.96 dB, is the characteristic 1-bit quantization loss and one of the main results from our analyses.

To look at the big picture for a moment, let us consider when different types of system solutions might be used in order to process a collection of samples. First, a mathematically ideal system which extracts all information of interest from the samples gives the absolutely best performance, but depending on what kind of operations that involves, this might not be implementable in a real-world system. Therefore, the second alternative, which is analog averaging, might be used instead. This will probably be a good or even the best possible solution in many cases. Then there is the third option, stochastic resonance. This is usually a sub-optimal solution, but it might still be of great interest because 1-bit logic in modern technology can be cheaper and more feature-rich than AA. However, as shown above, there might be some special cases where the SSR solution could actually be better than AA, i.e. be closer to the mathematically ideal solution than AA is. The trident noise distribution which demonstrates this is somewhat contrived, but it is still possible that there are some actual cases where the input noise is such that SSR is, in fact, better than AA.

Similar results to those shown above have already been reported in previous work. An SNR analysis resembling ours has been done, and it yielded a 2 dB loss in the output SNR for Gaussian noise and a 7 dB loss for sinusoidal noise [Reml 66 ☺], just as the results in table 4.1. The SNR loss due to quantization has also been analyzed using a different method than here, again yielding the same result, 1.96 dB loss for 1-bit quantizer, but also pointing out that a 3-bit quantizer only has a 0.265 dB loss [Zeng 13 ☺]. Note that even in GPS applications, where the SNR is quite low, the 1.96 dB quantization loss figure is not actually reached in receivers using 1-bit quantization. This is because the SNR just is not low enough to reach the “low input SNR” condition. The figures are instead 3.5 dB or 2.25 dB depending on Intermediate Frequency (IF) bandwidth [Braa 99 ☺].

When it comes to our demonstration of SSR being better than AA in some cases, that has also been established before. It has been shown that SSR occurs also for impulsive or infinite-variance noise and that SSR is robust against the violent noise impulses or outliers that come with such noise distributions [Kosk 01 ☺, Kosk 04 ☺]. This kind of noise can also be called heavy-tailed and leptokurtic [W]. The noise type can be encountered in e.g. ocean acoustics, and it was shown that SSR could be used as a preprocessor for Direction Of Arrival (DOA) estimation in sensor arrays and that it would give significant SNR gains and significant improvement in DOA estimation under these noise conditions [Roy 06 ☺]. So there might actually be real cases where SSR outperforms AA. It has also been shown that SSR can be significantly better than a linear matched filter under these conditions [Hari 09 ☺].

4.11 Large-Signal Analysis

Figure 4.18 shows the output error characteristics of an SSR system with Gaussian input noise. The output PMF shows the probabilities for the two quantizer output values, plotted against input signal value. When the input signal is 0, the probabilities are both 0.5, and for other values, the two probability values shift accordingly. The expected output value is the mean of the PMF, and since the input noise CDF is not a straight line, the expected-value curve will not be straight either. This will add distortion error to the output. Due to the random switching between the two quantizer output levels, there are random fluctuations with a mean of zero around the expected output value. We will call this random error. It is usually the total output error which is of practical interest, though, i.e. the output PMF compared to the input signal value. It is possible to compensate for the distortion to some degree, but that requires a more complex system and will not be considered here.

The RMS amplitudes of the distortion error and the random error can be RMS-added together to find the total output error RMS amplitude, known as Root-Mean-Square Error (RMSE). This decomposition and recombination can be done because the two errors are uncorrelated. To show that this holds, assume we have an input signal x , output distortion error y , and output random error z so that the total output signal is $x + y + z$. The mean-square value of the output error $y + z$ is then $E[(y + z)^2] = E[y^2 + 2yz + z^2] = E[y^2] + E[2yz] + E[z^2]$. Since y and z are uncorrelated, the middle term will be zero and we get $E[(y + z)^2] = E[y^2] + E[z^2]$. This summing is shown in the third row of the figure.

4.11.1 With an Input Signal Distribution

If the input noise strength σ_N is fixed and the input signal strength increases, the output error will decrease as more of the input signal is placed into the “W” valleys in the curve for the total output error. As the input signal strength continues to increase, the distortion error, i.e. the outer sides of the “W”, will take over and the output error will increase again. A moderate amount of input noise will therefore give less SNR loss than the 1.96 dB SNR loss we get when there is much noise. This can be seen in the plot of output SNR vs. input noise in figure 4.23, where the SNR metric A curve for $n = 1$ has 1.96 dB SNR loss for large input noise and much less loss when the input signal and the input noise have about the same strength. The same can be seen in the plot of a spectrum simulation with a noise sweep in figure 4.49. Note that this is something more than just the SR effect, i.e. in addition to exhibiting an SR peak, the system also performs closer to an ideal analog system under certain conditions. Also note that all these results are for SNR metric A, as described in section 4.9, with quantizer output levels set according to the small-signal analysis in section 4.10. As we can see in figure 4.23, SNR metrics B and C do not exhibit the kind of behavior described here. This tells us that it is important to keep in mind that the choice of SNR metric determines which system properties we will observe. The peculiarities of the SNR metric are discussed further in section 4.9.9.

We can explain the reduction in output error, i.e. output noise, in a different way too. The quantizer has output levels of e.g. ± 1 , which means that the total output power will be constant. Since the total output is the output signal plus the output noise, an increase in output signal

power could cause a reduction in output noise power. This explains the noise reduction we might see when the signal power increases. Note that such a reduction does not necessarily happen since the output signal and the output noise could be correlated so that their powers do not add linearly. As an example, an input signal value of 10 could cause an output value of 1 and thus an output error value of -9 . In this case the output noise power would in fact increase as the signal power increases.

Even though it might seem like a good idea at first, we can not bias a signal into a valley in order to reduce the SNR loss. This is because the signal gets attenuated more than the noise there, which will actually increase the small-signal SNR loss in the valleys. The small-signal SNR loss is plotted in the bottom row.

An example of where moderate input signal values will cause a reduction in RMSE at the output, is in the lower right image in figure 4.9, where the light and dark areas will exhibit a lower RMSE than the 50 %-gray areas. In other words, there will be less than 1.96 dB SNR loss in those areas. This observation could come as a surprise if one does not know about the effect described here.

4.11.2 Integration

By using integration, i.e. by averaging several results from the stochastic quantization, we can get a less noisy output signal. As can be seen in the third row of the plot, the integration only reduces the random error, not the distortion error. With much integration, there is less of a “W” shape in the total output error and the distortion error becomes dominant for smaller input signal amplitudes than with little integration.

4.11.3 Uniform Noise

Figure 4.19 shows corresponding plots for uniform input noise. An interesting observation to make here is that the small-signal SNR loss actually gets lower when the input signal value moves away from zero, and for input signal values close to ± 1 , there is even improvement instead of loss in the SNR. Note that it is just the SNR value which increases when the signal goes through the system in these cases, not e.g. the information content. The SNR metric sometimes behaves like this even though it might at first sound impossible. In other words, one must be careful when interpreting SNR values. This is discussed further in section 4.9.9.

To understand the SNR increase we can sometimes get with uniform noise, it might be illuminating to look at binomial distribution PMFs corresponding to integration of quantizer output values, shown in figure 4.20. We do $n = 100$ trials, i.e. integrations, with two different probabilities p , i.e. different input signal values. The PMF shows the probabilities for each possible total count. A binomial distribution has $\mu = np$ and $\sigma = \sqrt{np(1-p)}$, so the mean count value is proportional to p and thus a linear function of the input signal value for uniform input noise. We see that a low p , i.e. an input signal value close to -1 , gives low spread in the PMF, i.e. low noise, and that $p = 0.5$, i.e. an input signal value of zero, gives more spread. The closer the input signal value is to -1 , the lower the noise will be, just as shown in the plots for output error and small-signal SNR loss.

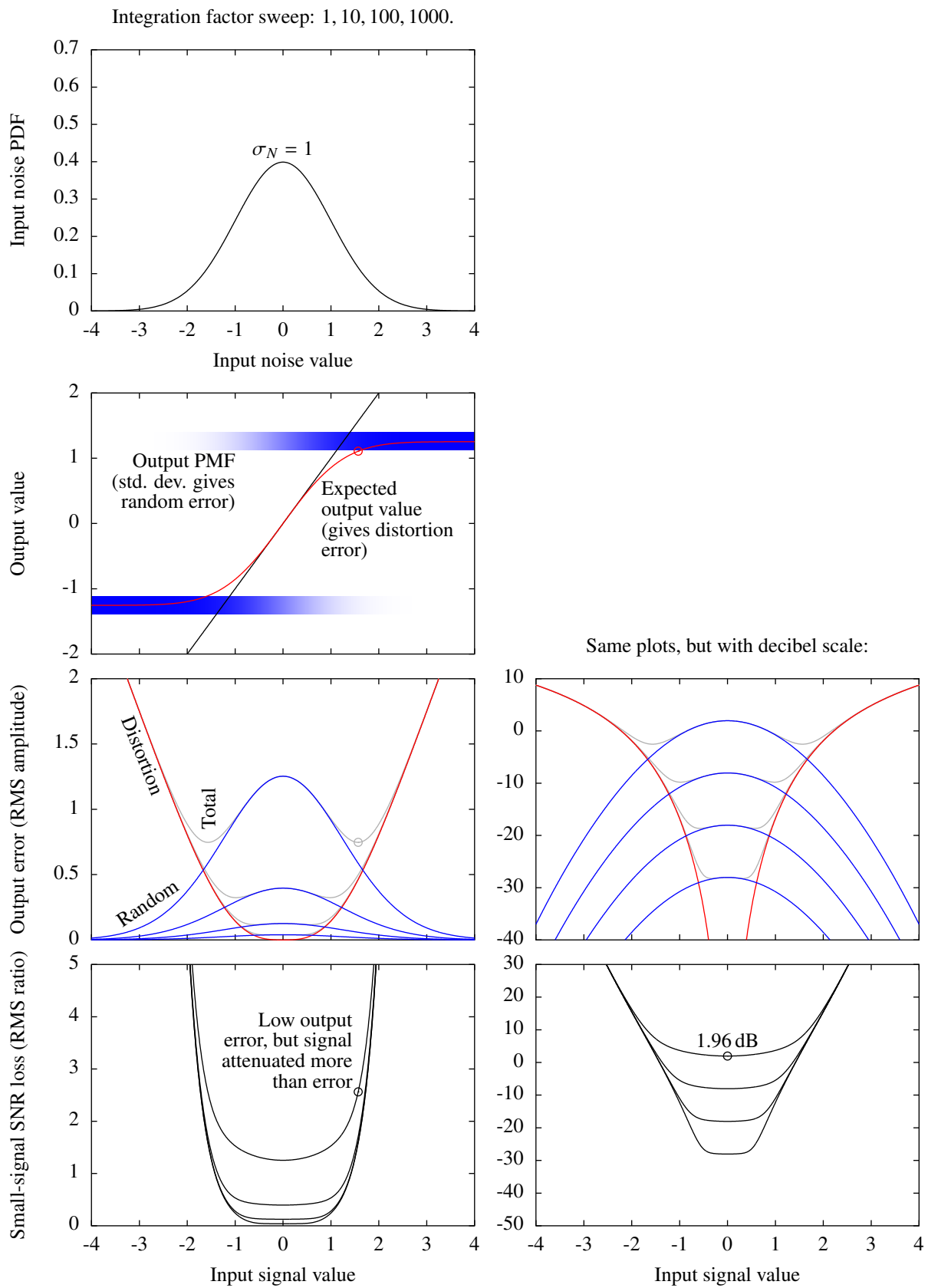


Figure 4.18: Decomposed output error and small-signal SNR loss of an SSR system with Gaussian input noise.

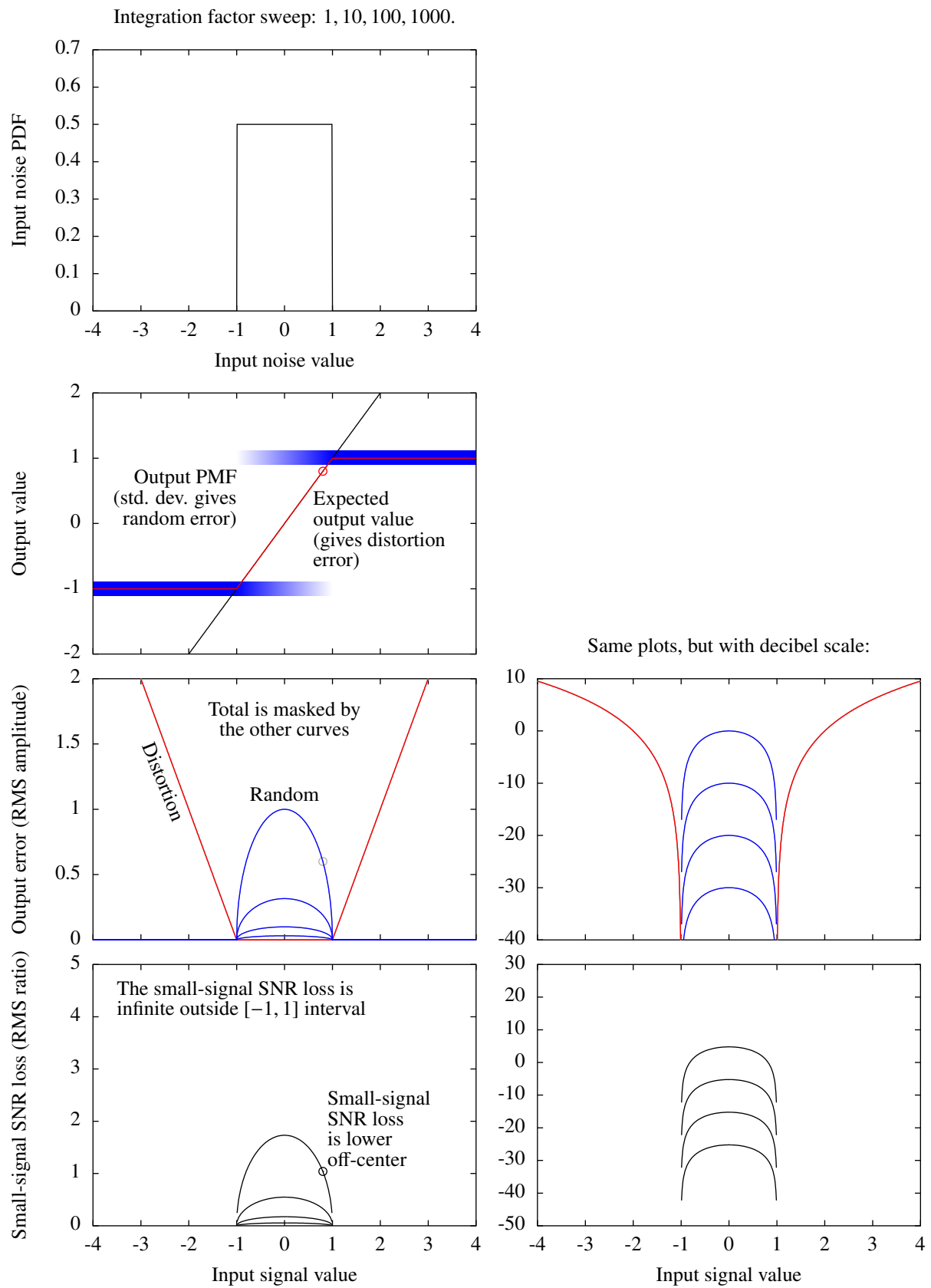


Figure 4.19: Decomposed output error and small-signal SNR loss of an SSR system with uniform input noise.

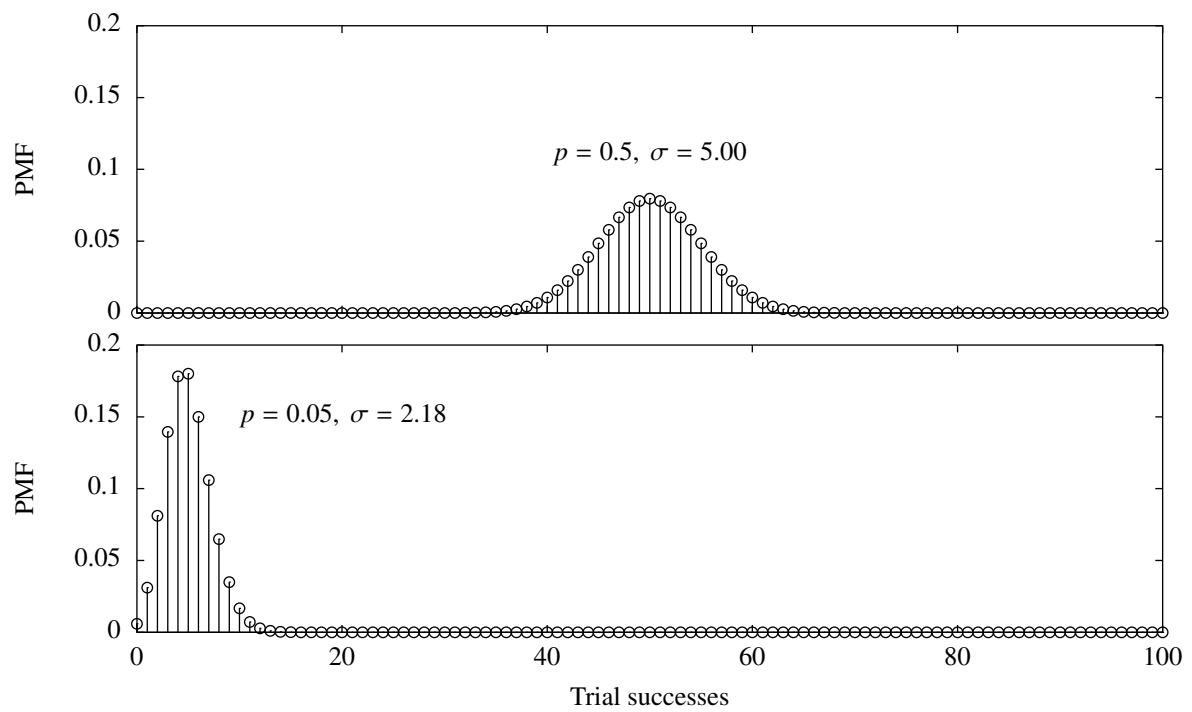


Figure 4.20: Binomial distributions corresponding to integration of quantizer output values.

4.12 SNR As a Function of Input Noise

Figure 4.21 shows the SNR at the output of a 1-bit quantizer plotted against the RMS amplitude of the input noise. The input to the system is a white Gaussian signal with white Gaussian noise, where the RMS amplitude of the signal is kept at $\sigma_{\text{signal}} = 1$ and the RMS amplitude of the noise σ_N is swept. The system is a simple 1-bit quantizer with a threshold of 0 and quantizer output levels which could be e.g. ± 1 . Signal integration is done by parallelization, not oversampling, i.e. the same signal but independent noise is sent through many 1-bit quantizers and then averaged. There is no distortion compensation. Since both signal and noise are white and we do not oversample, the samples will be independent and we therefore do not need to consider the time dimension. This means that all we have to do is to analyze the probability distributions associated with a single sample, which we do using numerical methods. The SNR value is calculated using SNR metric C, as described in section 4.9.

Figure 4.22 shows a decomposition of the noise part of the SNR. The two components are random noise and distortion error.

From the top plot, we can see that more integration gives higher SNR, at least when there is much input noise. This is because the standard deviation of the random noise gets divided by \sqrt{n} when we average the results from n 1-bit quantizers, as can be seen in the bottom plot.

When there is no noise, the integration just averages multiple copies of the exact same value and therefore has no effect. The SNR will consequently stay at about 2.5 dB regardless of how much integration is applied.

When there is a moderate amount of noise, the DC transfer function from the input signal to the expected output value has distortion so that even without the random noise there would still be much output noise due to the distortion error, as can be seen in the bottom plot. This can also be seen in the top plot, where the SNR limit imposed by the distortion error is plotted. The limit curve corresponds to infinite integration, i.e. no random noise in the output. We can clearly see how the behavior of the 1-bit quantizer is close to that of an ideal analog system when there is much noise, but that the distortion error limit pushes the curves down, making the characteristic SR peaks. The near-ideal behavior is analyzed in section 4.10 and the DC transfer function can be seen in figure 4.18, where we can see why it has a linear behavior for low-SNR inputs and why it gives distortion for high-SNR inputs.

To put it another way, when the input signal is white Gaussian and there is a moderate amount of input noise, the expected output value will be a compressed version of that input signal, close to a completely thresholded signal when there is very little noise. The random noise, which is the only noise component that gets reduced by integration, is on top of the compressed signal. By integrating a lot, the random noise could be reduced to negligible amounts, but the signal would still be compressed. The difference between the correct signal and this compressed signal is what we call the distortion error, and it is included in the noise part of the SNR. The output noise will therefore be large when there is much compression and the SNR will thus be low, no matter how much we integrate.

We can see the SR effect, i.e. SNR improvement by adding noise, when we use integration, i.e. for $n > 1$. Without integration, adding noise is just detrimental. The bottom plot shows that when we have the optimal amount of input noise, i.e. when the SNR is at its highest, reducing

the input noise would give us more distortion error and increasing the input noise would give us more random noise, in both cases reducing the SNR. By designing a system to have an optimal input SNR, though, we can maximize its performance.

With large amounts of input noise, i.e. with a low input SNR, we see that the system will have an SNR loss of only 1.96 dB compared to an ideal analog system. Since the input signal is so small compared to the input noise in this case, we can perform a small-signal analysis of the 1-bit quantizer. This is exactly what is done in section 4.10, where it is shown analytically that the SNR loss is 1.96 dB.

4.12.1 SNR Metrics A, B, C

Figure 4.23 is the same kind of plot as figure 4.21, except that all three SNR metrics A, B, and C are shown instead of just C. The main property of each of the SNR metrics are listed in the figure. Further description can be found in section 4.9. For SNR metric A, quantizer output levels of $\pm \sqrt{\pi/2} \cdot \sigma_N$ are used, as given by the small-signal analysis in section 4.10.

The three SNR metrics share some characteristics and are even equal in some regions, but they also yield quite different results in other regions. This shows that it is important to choose the right SNR metric. SNR metric B always yields the highest SNR value, and is always 0 dB or higher. SNR metric A coarsely matches SNR metric C and is identical to it for large amounts of input noise. This means that SNR metric A can be used as a much simpler implementation of SNR metric C if the input SNR is low enough. This is exploited in the spectrum simulations in section 4.14.

One of the spectrum simulations, figure 4.52, shows a sweep of the input noise just like the plots here, except that the spectrum is plotted instead of the SNR. The simulation uses SNR metric A with quantizer output levels set according to the small-signal analysis, just like here, but unlike here, integration is done by oversampling instead of parallelization. The characteristic property that a moderate amount of input noise is optimal can be seen in that plot too, though.

Another thing we can see in figure 4.23, is that the curve for SNR metric A with $n = 1$ has a particularly low SNR loss when the input signal and the input noise have about the same strength. This is discussed in section 4.11.

Figures 4.25 and 4.26 show the decomposition of the noise for SNR metrics A and B.

4.12.2 Quantizer Output Levels

Figure 4.24 shows what the quantizer output levels must be in order to give the signal part of the output the same amplitude as the input signal. What constitutes the signal part is defined by the given SNR metric. Note that only one of the two quantizer output levels is plotted, and that the other output level must be set to the same value as the first one but with opposite sign. As we can see, the curves are quite different, so also here it is important to choose the right SNR metric.

SNR metric A does not actually decide itself what the output levels should be since its job is to simply measures the performance of the system with whichever output levels have been chosen. Here we use output levels $\pm \sqrt{\pi/2} \cdot \sigma_N$, as given by the small-signal analysis in section 4.10. This gives correct behavior for large amounts of input noise but is not really intended for low amounts of noise. As we can see in the plot, the output level curve is identical to that of SNR

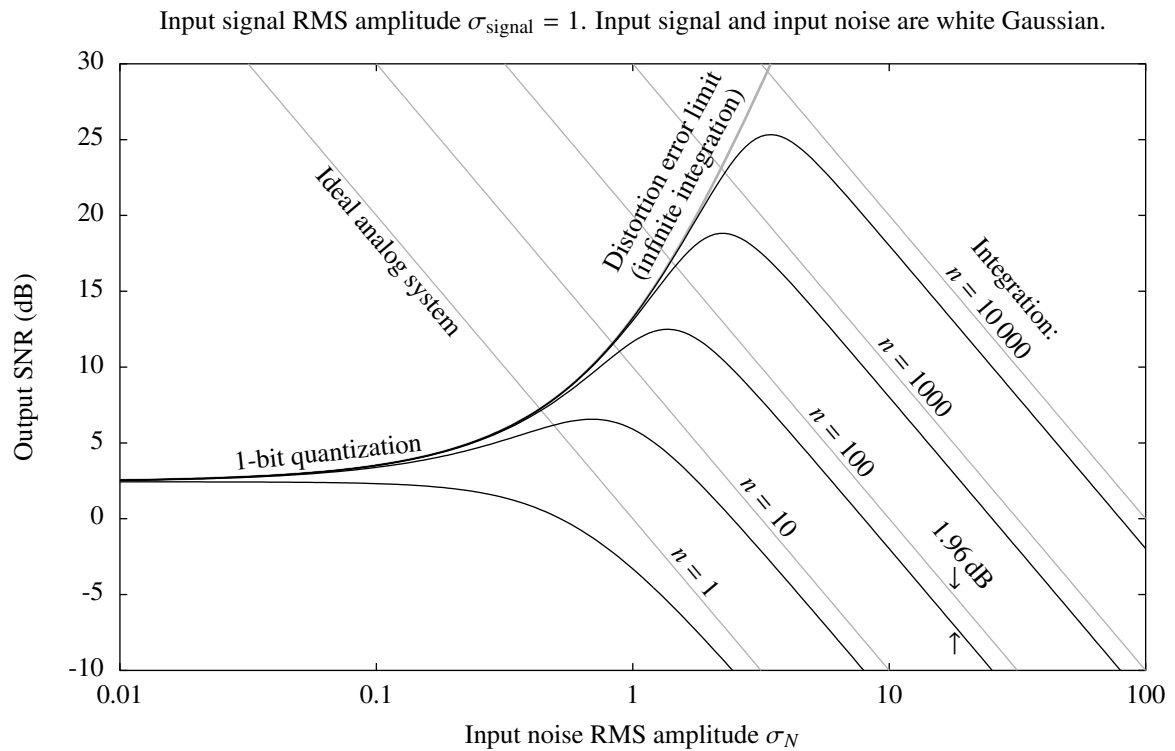


Figure 4.21: Output SNR vs. input noise for a 1-bit quantizer, i.e. an SSR system. Using SNR metric C.

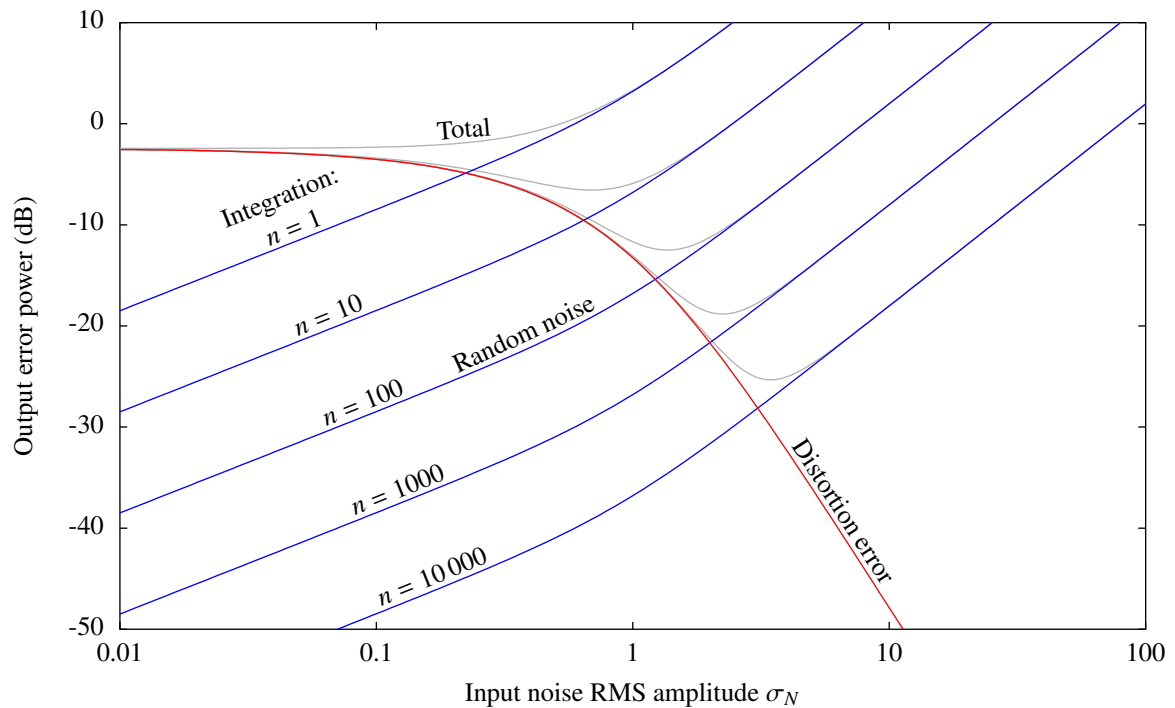


Figure 4.22: Decomposition of the plot above.

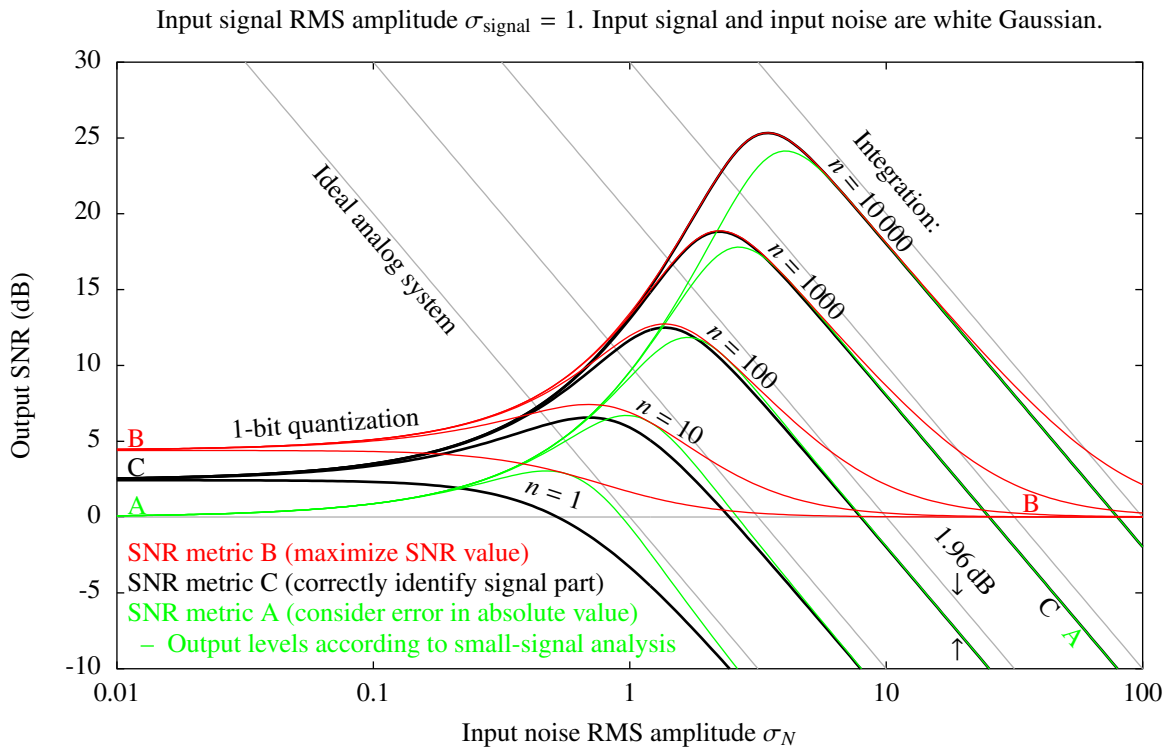


Figure 4.23: Output SNR vs. input noise for a 1-bit quantizer, i.e. an SSR system. Using SNR metrics A, B, and C.

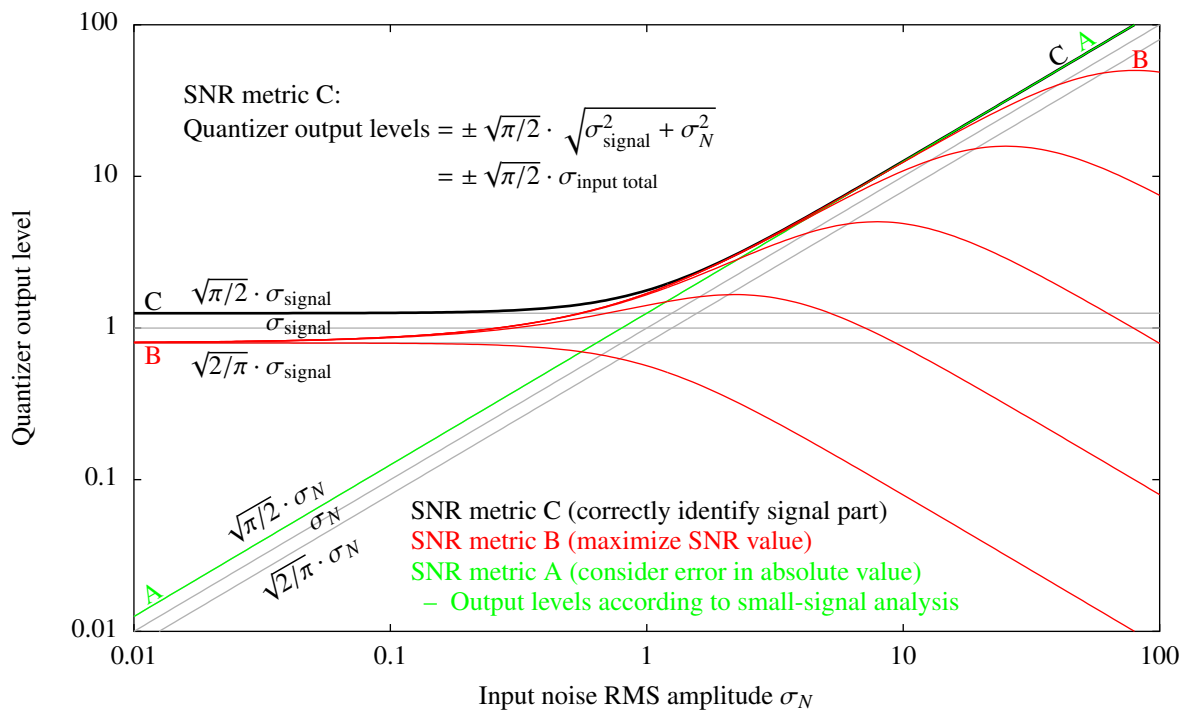


Figure 4.24: Quantizer output levels for the system in the figure above which would give the output signal, as defined by the given SNR metric, the same amplitude as the input signal.

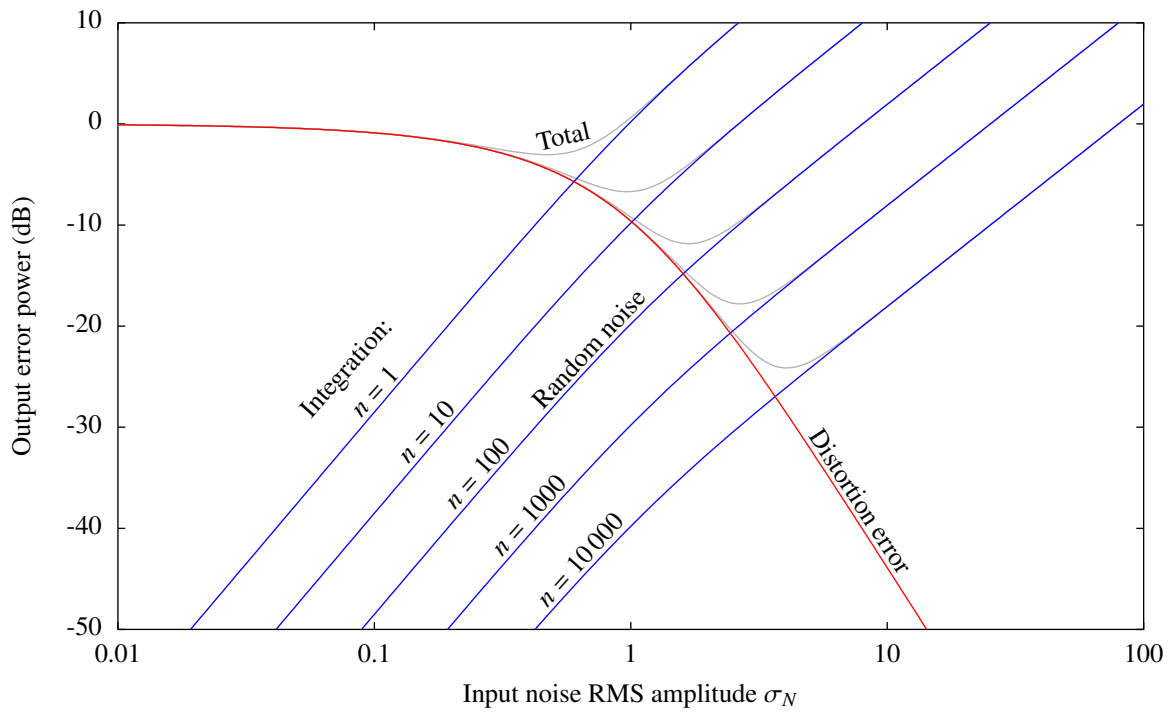


Figure 4.25: Decomposition of the curves for SNR metric A in figure 4.23.

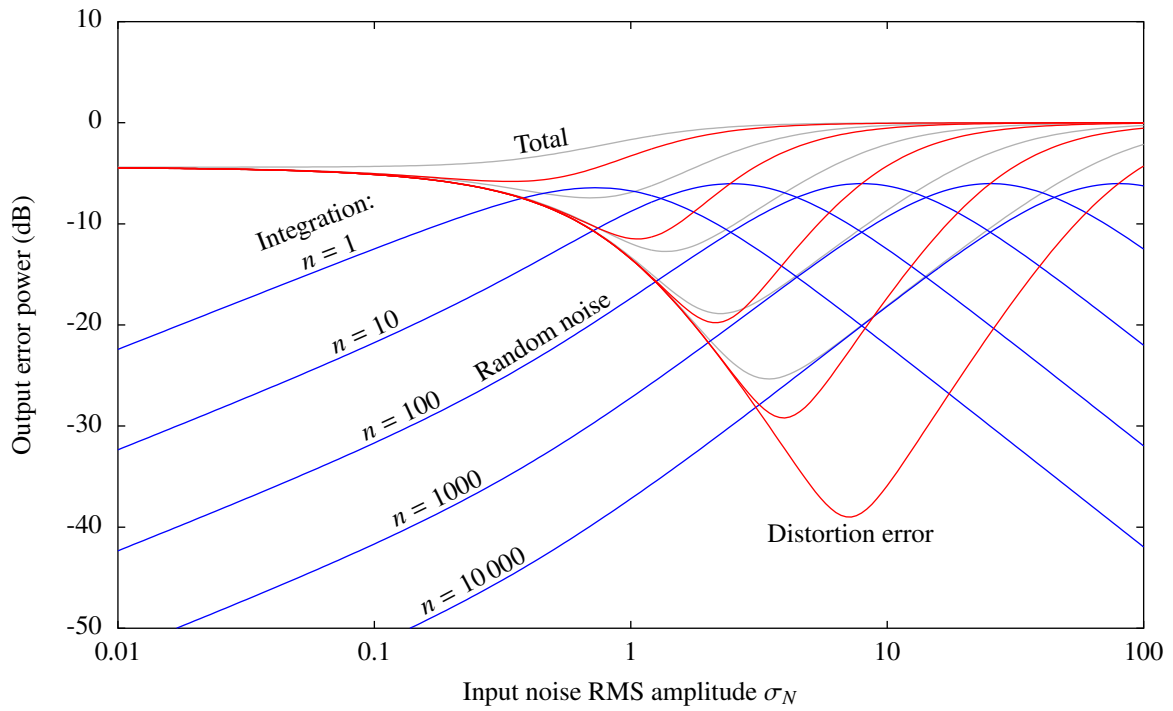


Figure 4.26: Decomposition of the curves for SNR metric B in figure 4.23.

metric C for large amounts of noise. As mentioned above, this means SNR metric A can be used as a simple replacement for SNR metric C under such conditions. In other words, if we know that the input SNR is sufficiently low and we want to measure the SNR using SNR metric C, we can just set the output levels to $\pm \sqrt{\pi/2} \cdot \sigma_N$ and do a naive SNR metric A measurement instead of performing the more complex calculations required for SNR metric C. In fact, SNR metric A can mimic any of the two SNR metrics under any noise conditions as long as the output levels are set to the correct values. These values can be read from the plot.

SNR metric B gives output levels that do not follow a simple pattern and that even vary with degree of integration. We can see that the more random noise there is at the output, i.e. the more input noise and the less integration there is, the more SNR metric B reduces the quantizer output levels compared to the other metrics in order to minimize the MSE at the output. This maximizes the SNR value, which is the purpose of SNR metric B.

SNR metric C gives output levels $\pm \sqrt{\pi/2} \cdot \sqrt{\sigma_{\text{signal}}^2 + \sigma_N^2} = \pm \sqrt{\pi/2} \cdot \sigma_{\text{input total}}$. This expression was just inferred from the numerical results in the plot, but it seems to fit the curve precisely. A possible explanation for the expression is that if we only consider a very small component of the original input signal as the signal and consider the remaining part of the input signal and the original noise as the noise, then this new noise will have $\sigma_{N \text{ new}} \approx \sqrt{\sigma_{\text{signal}}^2 + \sigma_N^2} = \sigma_{\text{input total}}$. By then using the expression from the small-signal analysis, we get quantizer output levels $\pm \sqrt{\pi/2} \cdot \sigma_{N \text{ new}}$, i.e. exactly the expression for SNR metric C. The small component of the signal would then, according to the small-signal analysis, have the same amplitude in the output as in the input.

The quantizer output levels from the plot can be used when designing a system where the output signal should have the same amplitude as the input signal. One example of this is if we are interested in the absolute value at the output and we want to minimize the MSE there. We would then use the output levels given by SNR metric B for the amount of input noise and integration we have. A specific case is when there is negligible or no input noise. The output levels must then be $\pm \sqrt{2/\pi} \cdot \sigma_{\text{signal}}$. Another example is if there is a lot of input noise and we want the expected output value to be equal to the input signal. We would then use the output levels which we have here chosen for SNR metric A, namely $\pm \sqrt{\pi/2} \cdot \sigma_N$ from the small-signal analysis. As discussed above, SNR metric A can also be used as a simple substitute for SNR metric C in this case. Such a substitution is done in the simulations in section 4.14.

4.12.3 Related Work

To use quantizer output levels as given by SNR metric B is also known as Wiener decoding [McDo 05] [McDo 06b, pp. 199]. Such Wiener decoding has been used in literature to produce plots similar to some of those in figure 4.21–4.26 for SNR metric B. An SNR plot has been made [McDo 05, fig. 4] and has also been compared to the mutual information metric. One interesting observation from this is that the two metrics have their SR peaks at different input noise strengths. The output error power has been plotted too [McDo 06b, fig. 6.7(a)], but only the total error power, not its decomposition. Lastly, the quantizer output levels, or reconstruction points for Wiener decoding, have also been plotted [McDo 06b, fig. 6.8(a)]. Note that some of these plots are not logarithmic and that they all have a limited input noise range, so they may

look a bit different from those presented here. In addition to these plots for SNR metric B, an SNR plot for something close to SNR metric C, which seems to just ignore the distortion error component, has been made too [Roy 06 ♻].

4.13 Colored Noise

A common assumption when analyzing oversampled SSR systems, is that the input noise is white, i.e. that its spectrum is flat and that the noise samples are uncorrelated. This means that there could be a brick-wall anti-aliasing filter in front of the sampled SSR system, so that the wideband white input noise gets cut cleanly at $f_s/2$. With a less sharp anti-aliasing filter, the spectrum of the sampled signal might not be completely flat.

Instead of using white noise like this, we could use colored input noise, such as first-order low-pass filtered Gaussian noise. This noise filter could be put in front of the anti-aliasing filter or could simply replace it. In the latter case, aliasing issues should be kept in mind. Because the noise might extend further up in frequency and because the 1-bit quantizer also spreads the noise to higher frequencies, it is often assumed that the SSR system is continuous-time or has a high sampling rate compared to the upper cut-off frequency of the noise filter. In other words, we could keep the 1-bit quantizer but remove the sampler. The SSR system could then look like this: Wideband white Gaussian noise \rightarrow noise filter \rightarrow continuous-time 1-bit quantizer. With a system set up like this, it is interesting to see what happens to the noise spectrum in the output. As it turns out, we will be able to reduce the usual SNR loss of 1.96 dB to much lower values.

Colored input noise for SSR systems is also known as pre-filter, post-filter, oversampling (while keeping the original anti-aliasing filter), continuous-time system, unsampled, and dependent samples.

As shown in section 4.10, the usual SNR loss of an SSR system is 1.96 dB. For example first-order low-pass or brick-wall filtered noise gives less SNR loss [Beau 86 ⑤, ch. III] [Beau 86 ⑤, p. 45]. Firstly, by using a first-order low-pass filter for the noise, we get less than 0.5 dB SNR loss. About 10 times oversampling is here sufficient to give a behavior close to a continuous-time system. By this, we mean that $f_s/2 = 10 \cdot f_c$, where f_s is the sampling rate and f_c is the cut-off frequency of the filter. Secondly, by using a brick-wall filter, we get about 1 dB SNR loss. About 3 times oversampling is sufficient here. Similar results have been found for PCA detectors [Kane 66 ⑤], which are systems that are somewhat similar to SSR. The spectrum simulations in section 4.14 yield similar results too, as can be seen in the spectrum plots in figure 4.30 and 4.31.

For narrow-band signal and noise, it has been found that the SNR loss due to quantization only, i.e. without sampling, is $4/\pi = 1.05$ dB, that sampling reduces the SNR further by $\pi^2/8 = 0.91$ dB, and that the total SNR loss is $\pi/2 = 1.96$ dB [Reml 66 ⑤]. This is confirmed by the single-frequency spectrum simulation in figure 4.37, but the narrow-band spectrum simulation in figure 4.35 does not agree. It is therefore important to note that these results are only valid for the very narrow-band case.

An interesting observation about SSR systems is that the quantization and the sampling each are responsible for part of the SNR loss [Kane 66 ⑤], as explained in the following. Continuous-time 1-bit quantization spreads the spectrum of the input, so the Nyquist rate of the quantized signal is not the same as for the input. Quantizing causes loss of only some information, and subsequent sampling takes the loss all the way to 1.96 dB because of the down-aliased noise. We assume the noise is filtered with a brick-wall filter, i.e. we simply remove the sampling operation from an ideal 1-bit quantization and sampling system. A different way to look at this, is that we

get more precise zero-crossing timings when we remove the sampling operation. The spectrum simulation in figure 4.30 shows the noise spectrum at the output of this continuous-time 1-bit quantizer. We see that the SNR loss is only about 1 dB and that there is some noise at higher frequencies too. If sampling were introduced, this high-frequency noise would get aliased down so that the SNR loss would be 1.96 dB.

Is it really only the noise that gets spread outside the band of the brick-wall filter, or is there any signal information there too? In other words, could the aliasing in a sampled system be bringing back some signal information which will be lost in a continuous-time system? It is shown in section 4.13.1 that colored input noise, such as brick-wall-filtered noise with very high sampling rate, does not change the expected value at the output of an SSR system compared to when the input noise is white. This means that the continuous-time quantization does not spread the input signal out to other frequencies, i.e. all components found outside the band of the brick-wall filter is noise, and the aliasing does not bring back down any signal information. Note that we here assume a small-signal input signal. A large-signal input signal would get distorted and would create harmonics which might get aliased. The spectrum simulations in section 4.14 confirm the analytical results given here, since no distortion or signal reduction due to the lack of down-aliasing can be seen there.

4.13.1 Why SSR with Colored Noise Works

In section 4.10, the SSR system was analyzed, but it was assumed that the input noise was white. So what happens when the input noise is no longer white? Will the SSR system still work as expected?

We will now consider a sampled SSR system. A continuous-time system can be approximated with a sampled system with high sampling rate.

To make the colored noise, we simply filter white Gaussian noise. The white Gaussian noise has independent samples, each with a Gaussian distribution. The filtering can be seen as convolution in time domain, so each sample of the filtered input noise is thus a weighted sum of independent stochastic variables with Gaussian distributions. This means that they have Gaussian distributions, just like the white Gaussian noise, but the difference is that the samples are no longer independent.

Since each colored noise sample has a Gaussian distribution, the expected value at the output of the quantizer is correct for any given sample, at least when seen in isolation. The dependence of the samples warrants a closer look at what happens when samples are combined in further processing, though.

As an example, consider an SSR system with low-pass-filtered input noise and a low-pass filter at the output of the system to recover the input signal. If a given input noise sample is a high value, the next sample will probably be a high value too. This means that the low-pass-filter at the output, which basically takes a weighted average of adjacent samples, will often get samples with similar values as inputs for this averaging. How will this affect the output of the filter?

Luckily, when taking a weighted sum of stochastic variables, the expected value of the weighted sum is the weighted sum of the expected value of each stochastic variable, even if the stochastic variables are dependent, i.e. $E[a_1X_1 + a_2X_2 + \dots] = a_1 E[X_1] + a_2 E[X_2] + \dots$, where

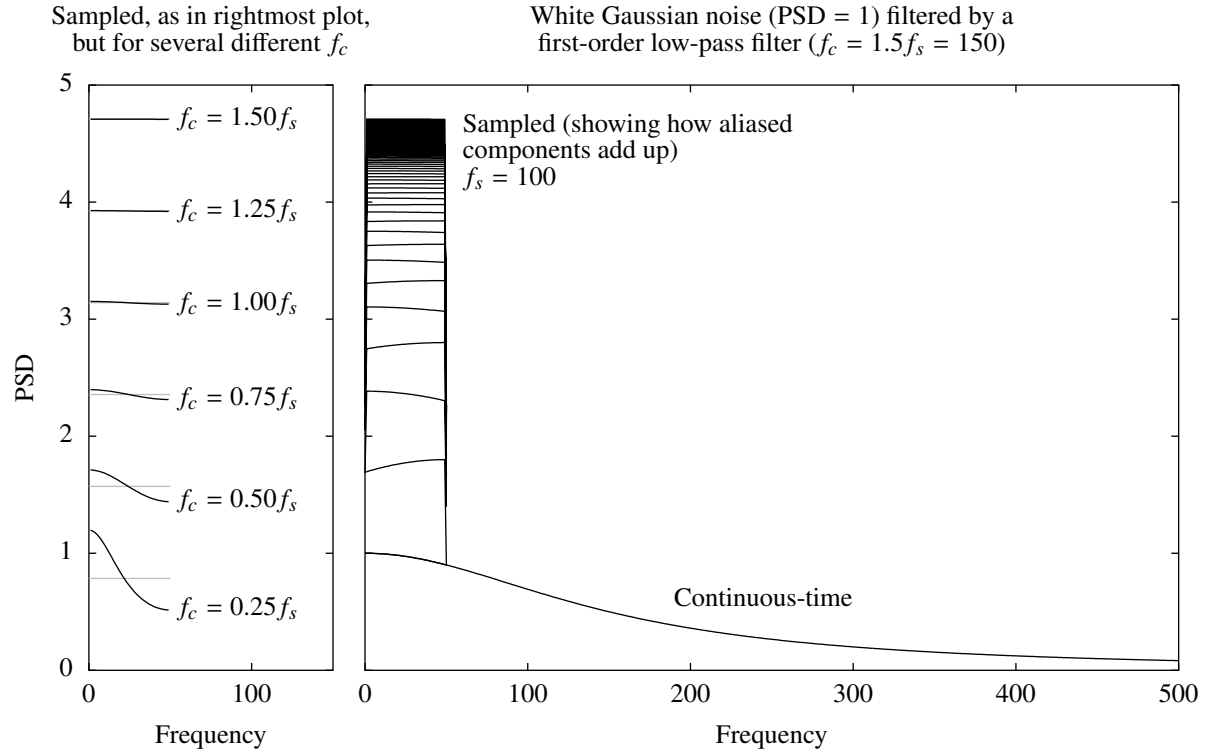


Figure 4.27: First-order low-pass filtered noise can be made white by sampling it. This is due to aliasing effects.

$E[\cdot]$ is the expected value operator, a_i are constants, and X_i are stochastic variables. This means that the output filter in the example will have the correct expected value at its output. In general, any further linear processing, such as filtering, will thus produce the correct expected value. In other words, SSR systems work well also with colored input noise.

Note that by “correct value” in the text above, we mean the input signal distorted by the input noise CDF curve. For small-signal inputs there is no distortion, so the “correct value” will simply be the input signal value, i.e. a completely linear processing of the signal.

The spectrum simulations in section 4.14 confirm the results of this analysis since no distortion due to colored input noise can be seen.

4.13.2 White Noise from First-Order Anti-Aliasing Filter

Figure 4.27 shows how white Gaussian noise filtered by a first-order low-pass filter and then sampled without an anti-aliasing filter, becomes white again if f_c is greater than about 1 to 1.5 times f_s , i.e. 2 to 3 times $f_s/2$. This can be used as a technique for creating white noise in the samples with the use of a simple first-order anti-aliasing filter instead of a sharper filter.

4.14 Spectrum Simulations

This section presents spectrum simulations of SSR systems, particularly demonstrating the effect of different input noise colors.

Figure 4.28 shows a typical equivalent block diagram for the simulations. The simulations are done as following. The input signal is in most of the cases single-frequency-filtered AWGN. The input noise is in most of the cases filtered AWGN. The signal and noise could be filtered together, but here we only filter the noise. When the filter, e.g. a first-order low-pass filter, has some attenuation in the signal band, the results might differ slightly since not only the noise, but also the signal, would be filtered if filtering were done on the sum of the signal and noise. A brick-wall anti-aliasing filter is used, also for e.g. the first-order low-pass case. Note that the quantizer and sampler can change places without affecting the result. The signal is 1-bit quantized with a threshold of 0. The quantizer output levels are $\pm \sqrt{\pi/2} \cdot \sigma_N$, where σ_N is the RMS amplitude of the input noise. “Continuous-time” is just a sufficiently high simulation sampling rate. The power spectra of many simulation runs are averaged in order to give smoother spectrum curves. Each simulation run uses a signal that is one time unit long, which gives a spectrum resolution of one frequency unit. We are using unit-less quantities. Note that these are simulations, so the results may not be exactly accurate. We may for example not exactly hit the 1.96 dB number. Note that while AWGN filtered to a single frequency will have a Gaussian distribution in the long run, a single simulation iteration of it will just be a sinusoid with a random amplitude and phase. This nuance could be important to be aware of in some cases.

The plots show spectra with PSD as the quantity. The curves that are plotted, are input signal, input noise, output without input signal, i.e. just quantized noise, which shows the behavior for low input signal amplitudes, and output error, i.e. output minus input signal, which includes both noise and distortion. Results such as SNR loss are shown at the top of the plots.

The plots do not show maximal achievable SNR conditions, so they can not be used to compare best achievable performance between different systems. The focus is on how noise is amplified, i.e. the SNR loss, and how it is moved in the spectrum when going through the system.

4.14.1 Measuring the SNR

We use SNR metric A, as defined in section 4.9, since it is simple, considers the absolute value, and can sometimes be used as a substitute for SNR metric C, as explained in section 4.9. The output levels of the 1-bit quantizer are set according to small-signal analysis to $\pm \sqrt{\pi/2} \cdot \sigma_N$, as shown in figure 4.6 and section 4.10.

The output error, as shown in the plots, is the “noise” part of our SNR calculations, so since we define it as the output minus the input signal, it contains both noise and distortion power.

Any distortion introduced by the system is thus considered noise. This is therefore a good SNR metric if it is the absolute value of the signal which is important, i.e. if any deviation in the output signal value from the unscaled input signal value is an error which should be taken into account.

When calculating the output SNR, we use the input signal as the signal part of the expression and the output error, i.e. the output minus the input signal, as the noise part. More precisely,

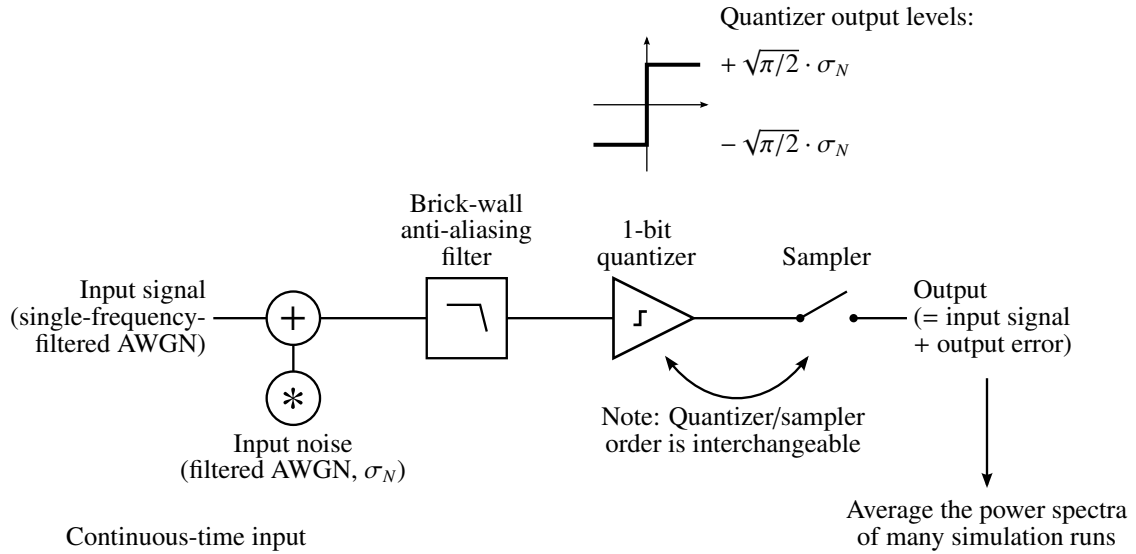


Figure 4.28: Typical equivalent block diagram for the 1-bit-quantization simulations.

we use the PSD levels, which are the values that are plotted in the figures, and we only use the values found within the input signal band. Note that some of the distortion power might fall outside the signal band and will thus be ignored. This might be different than what some SINAD definitions do.

In most cases here, the quantizer output levels are set to $\pm \sqrt{\pi/2} \cdot \sigma_N$, which for small-signal input signals will give no distortion and an output signal amplitude equal to the input signal amplitude. Even if the random noise at the output is large, the output levels are not reduced in order to increase the SNR as in SNR metric B. This means that the output levels are set according to SNR metric C and that SNR metric A, i.e. assumption-based, is then used subsequently. So when the input signal goes from small-signal to large-signal, we move away from SNR metric C and instead get the behavior of the more naive SNR metric A. The simplicity of SNR metric A can sometimes cause some of the output signal to be regarded as noise, such as in figure 4.52.

Note that SNR in general and SNR metrics other than C in particular can sometimes be deceptive and than e.g. information theoretic metrics can yield other results, as explained in section 4.9.

We mostly use the output error vs. input noise to calculate the SNR loss. In some cases, which are marked accordingly, we use the output when there is no input signal, i.e. just quantized noise, instead of the output error. This is the same as having no or an infinitesimal input signal and calculating the SNR as usual, using the output error. This is a more stable metric since it does not depend on the strength of the input signal.

4.14.2 Overview

The spectrum simulations that are done are as follows:

- Basic SSR
- Sampled vs. oversampled/continuous-time
- Filtered AWGN as input noise, i.e. colored noise:
 - Red (brick-wall (anti-aliasing filter), first-order, ideal attenuation (high-shelf))
 - Practical system with very low SNR loss
 - Green (band-pass) (with band-pass signal)
 - Blue (brick-wall, first-order)
 - Red/pink/white/blue/violet
 - Interesting discovery about flicker noise
- Half-tone / PWM / swept threshold (triangle)
- No noise (quantizer output levels $\pm \sqrt{2/\pi} \cdot \sigma_{\text{signal}}$)
- Sigma-delta for comparison
- Parameter sweeps to show the effects of distortion and of strong input signals reducing the output error power

4.14.3 Simulations

Figure 4.29: Basic SSR.

Figure 4.30: Brick-wall.

Figure 4.31: First-order low-pass.

Figure 4.32: High-shelf filters.

The first-order low-pass and high-shelves, but not the basic SSR, have the same full-band input SNR, i.e. the same $\sigma_{\text{filtered-input-noise}}$.

Figure 4.33 and 4.34: Practical system.

This is a more practical system setup. We do not use an ideal brick-wall anti-aliasing filter, but instead use a first-order filter for that purpose. Furthermore, we take into account the aliasing of the filtered AWGN components above the Nyquist frequency. We try to mimic the high-shelf filter, which has low SNR loss, but do so by using a simple second-order filter. In total it is thus a simple third order filter.

This setup is for when the input already has AWGN. If the input has very little noise to begin with, the noise shape used here might not be the optimal shape of noise to explicitly add (the low-frequency part of the noise just greatly reduces the SNR).

The frequency of the first pole of the filter, fp_1 , is chosen to be 31.6, so that in the frequency range from 0 to about 10 the filter response is close to unity. That way the input signal is minimally affected by the filtering. The frequencies of the zero f_z and second pole fp_2 are found by sweeping them and looking for the lowest SNR loss. An fp_2 of half a decade below the Nyquist frequency seems to be optimal. Any lower fp_2 and the power in the high-shelf, i.e. the rightmost part, would be reduced, which would reduce the important SSR effect it provides. Any higher fp_2 and the power above the Nyquist frequency would increase, and because of the lack

of an ideal anti-aliasing filter, the power would fold into the signal band and thus increase the SNR loss. The point of the ideal high-shelf filter is that the low-PSD and wide-band noise at high frequencies act as noise for SSR and thus only about 1.96 dB relative to that level is added at all frequencies. That makes the added noise at the low frequencies very low compared to the noise which is already there. So the high-shelf, mostly controlled by f_z , should be low-power in order to reduce the noise added in the signal band, but since the transition band is first-order here, the lower the PSD, the more the filter looks like just a first-order low-pass filter, and since such a filter gives as much as about 0.30 dB SNR loss, we must not come too close to the first-order filter shape if we want to achieve the very low SNR loss which a high-shelf filter can provide.

Note that the values found for f_z and fp_2 are close to optimal, but are not necessarily the exact optimal values, due to limited granularity in the sweeps.

The SNR loss metric we use here is how much higher the output without input signal, i.e. just quantized noise, is compared to the filtered but not sampled input noise. The other plots use the output error compared to the input noise. The reason for choosing this metric here is that adding the input signal changes the output noise, in fact so much that the SNR loss becomes negative in some cases. This phenomenon is discussed elsewhere. The SNR loss calculated without input signal yields the key result corresponding to e.g. the 1.96 dB of basic SSR, so it should be a good metric.

It should be possible to improve the SNR further by increasing the sample rate (and updating the filter frequencies). Or one could just increase the sample rate and pole 2, which should keep the SNR loss at about the same level, but increase σ_{noise} , thus allow wider input signal swing.

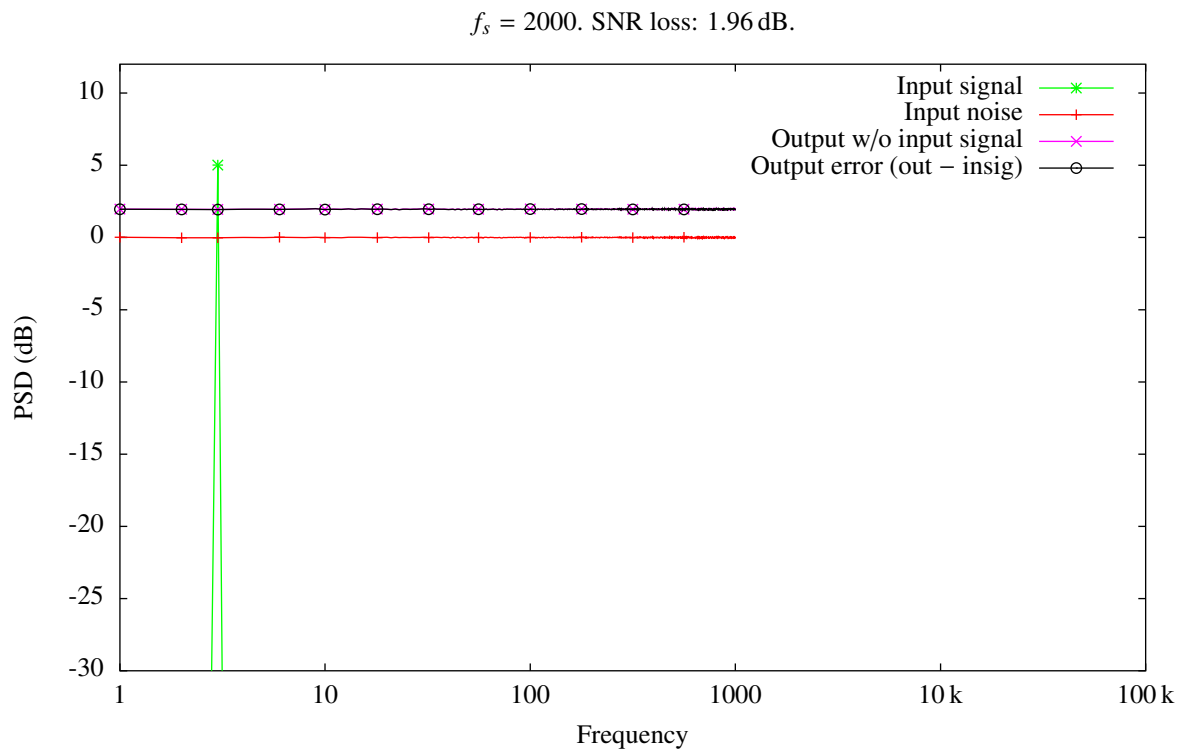


Figure 4.29: Spectrum of sampled 1-bit quantization with AWGN, i.e. a basic SSR system.

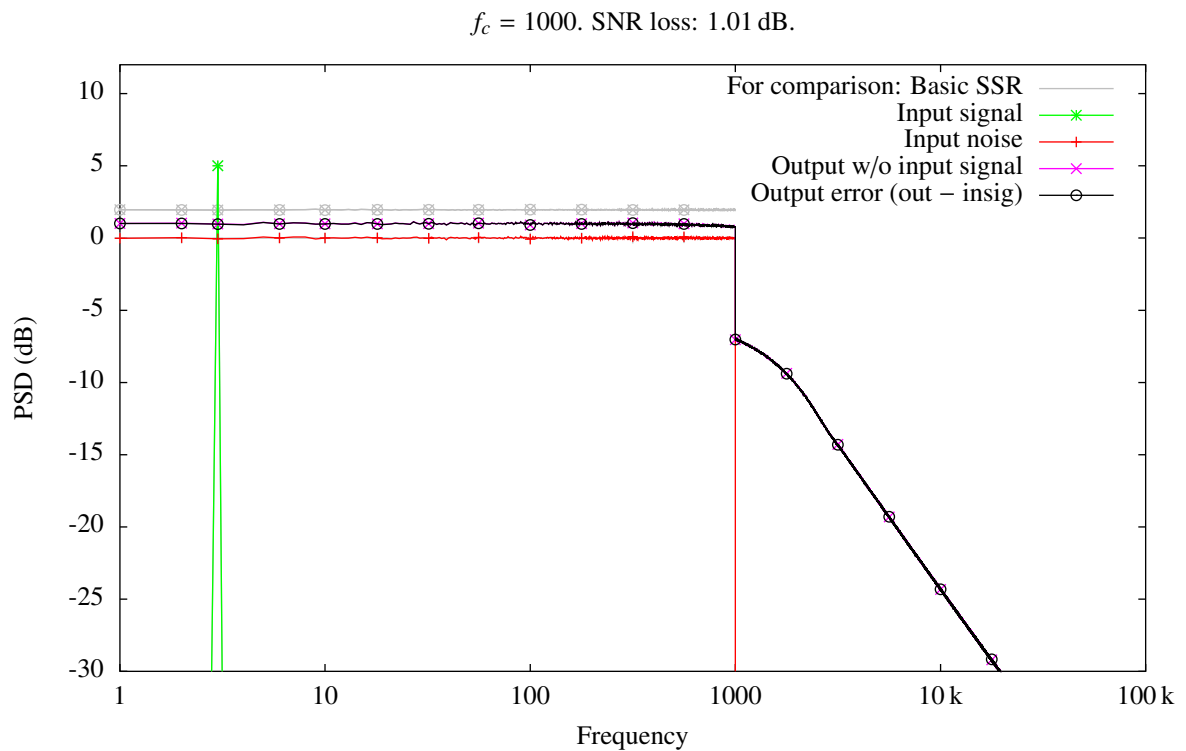


Figure 4.30: Spectrum of continuous-time 1-bit quantization with brick-wall-filtered AWGN, i.e. a basic SSR system with the sampling step removed.

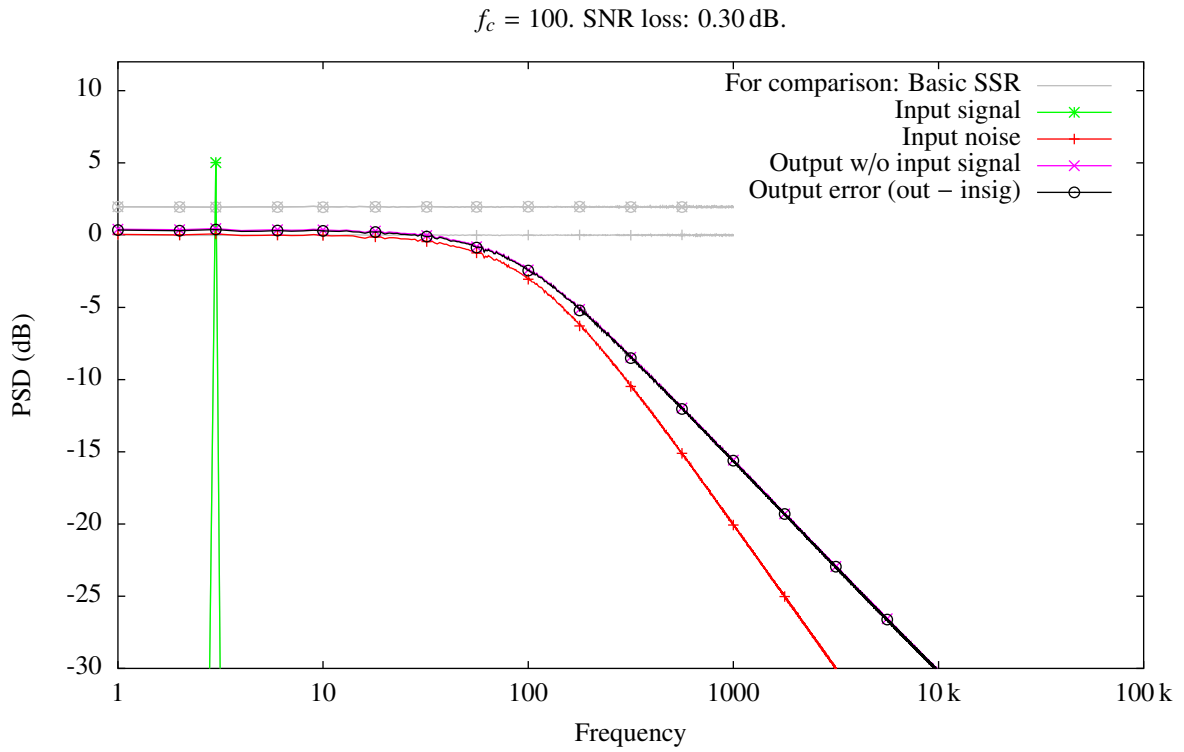


Figure 4.31: Spectrum of continuous-time 1-bit quantization with first-order-low-pass-filtered AWGN.

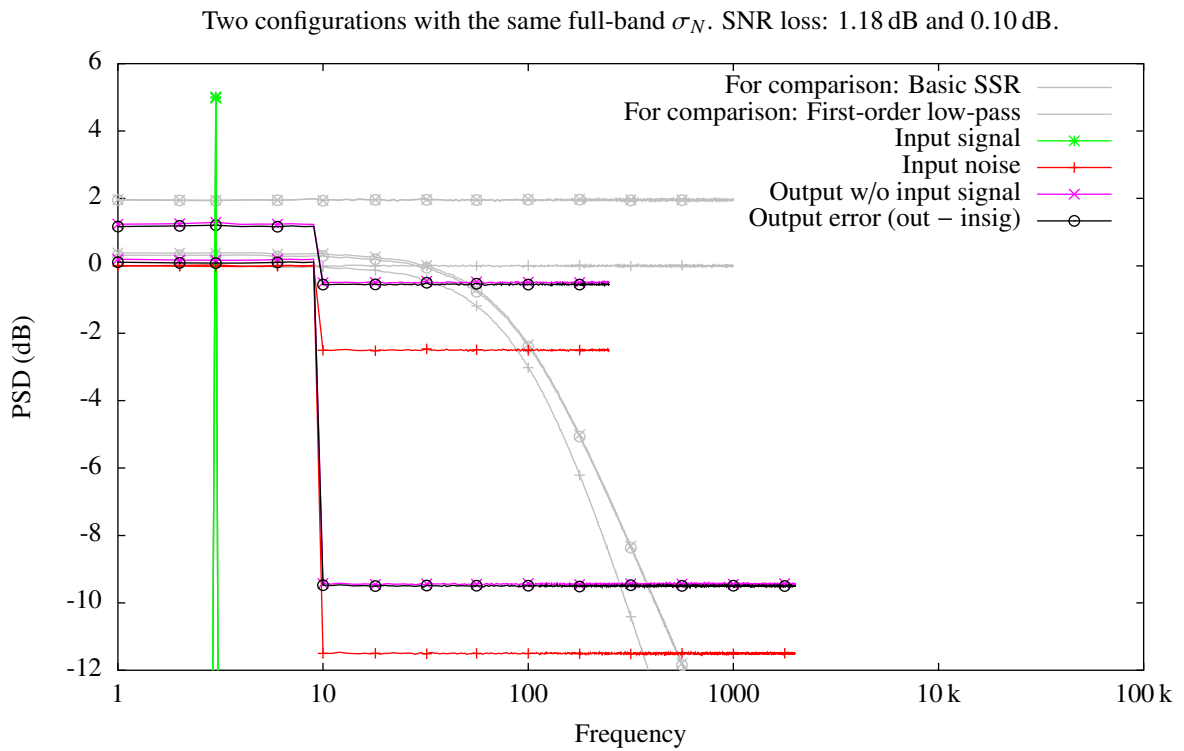


Figure 4.32: Spectrum of sampled 1-bit quantization with high-shelf-filtered AWGN.

$f_s = 20$ k. $f_{p1} = 31.6$. $f_z = 237$. $f_{p2} = 3.16$ k. SNR loss: 0.07 dB. SNR loss calculated w/o input signal.

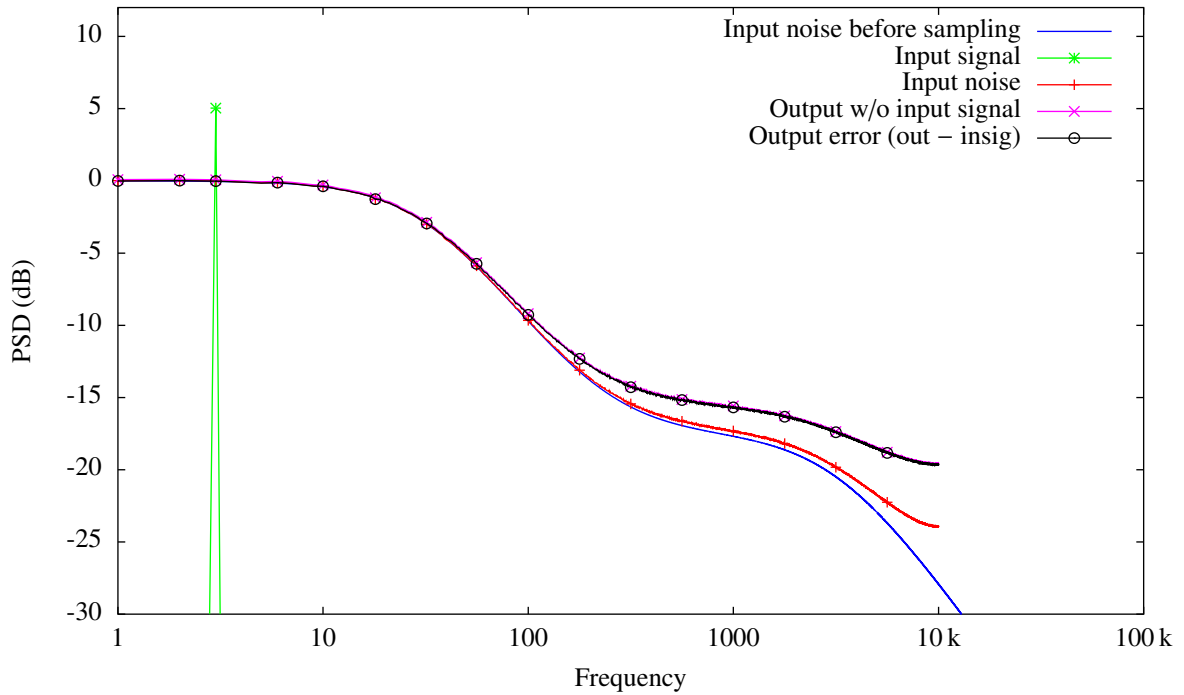


Figure 4.33: Spectrum of sampled 1-bit quantization with third-order-filtered AWGN and no other anti-aliasing filter than that, such as a brick-wall filter, i.e. a more practical system setup.

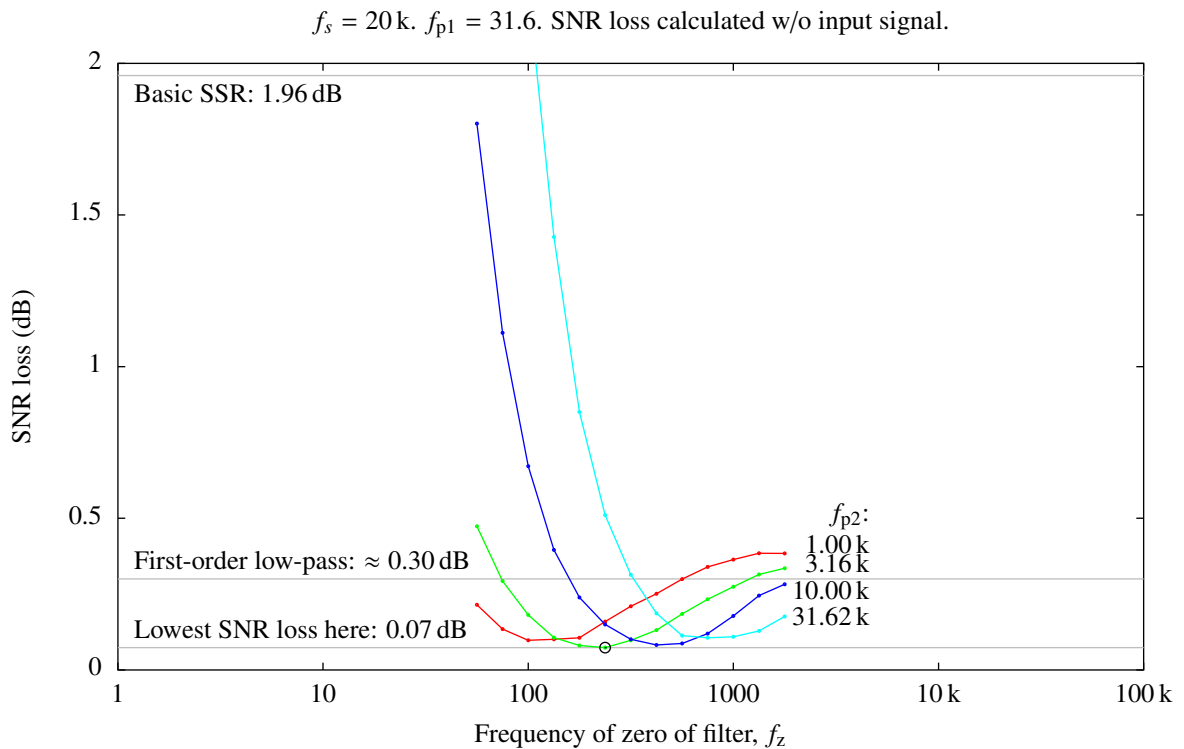


Figure 4.34: Parameter sweep for the system setup described in the figure above. The marked point is the configuration shown in that figure.

Figure 4.35 and 4.36: Band-pass.

Has peaks at odd harmonics. Signal can be outside noise band. About 0.6 dB SNR loss.

Figure 4.37: Single-frequency.

Single frequency. Has odd harmonics. Very low input SNR to prevent the input signal strength from affecting the result. This demonstrates the 1.05 dB loss described in [Reml 66 ⑤]. It seems only BW=1 works - any wider and the result deviates from 1.05 dB.

Figure 4.38: First-order high-pass.

Not sensitive to sample rate.

Figure 4.39 and 4.40: Brick-wall high-pass.

Sampled system. Changing the sample rate changes the result.

Figure 4.41: Power-law.

It seems -10 dB/decade (pink noise / flicker noise) has little or no SNR loss.

Figure 4.42 and 4.43: Flicker noise.

Flicker noise (-10 dB/decade). The larger the sampling rate is, the closer the SNR loss gets to 0 dB. Adding some AWGN seems to reduce the SNR loss even further. SNR loss calculated both with and without input signal present is plotted, and values for the latter case are shown at the top of the plot. As can be seen, and which is discussed elsewhere, having the input signal present can make the SNR loss vary by a great amount. Because of this, SNR loss calculated without input signal present is a more stable metric, and therefore it is these values that are shown.

What happens when the flicker noise extends to very low frequencies? The low-frequency noise can take the system out of its operating region a certain percentage of the time, which could be problematic in practice but which might not get caught by the averaged spectrum. The added white Gaussian noise might help here by increasing the amplitude of the high-frequency noise, thus making the operating region larger.

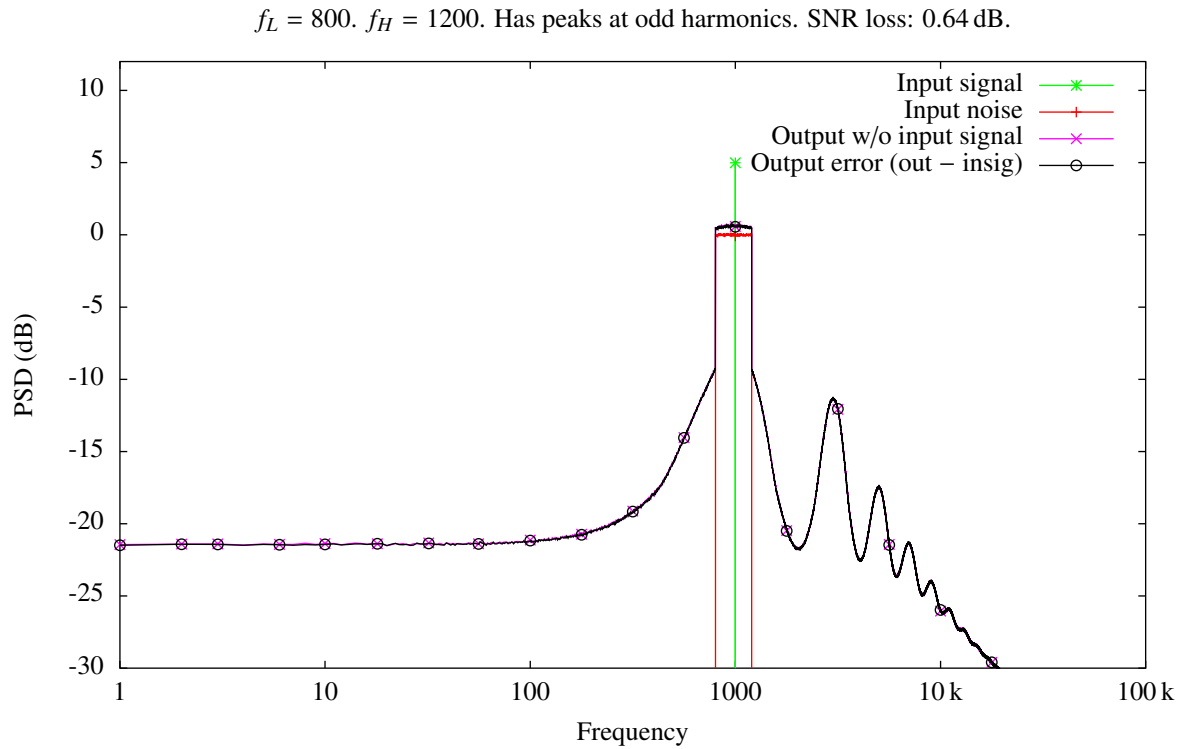


Figure 4.35: Spectrum of continuous-time 1-bit quantization with brick-wall-band-pass-filtered AWGN and signal inside the noise band.

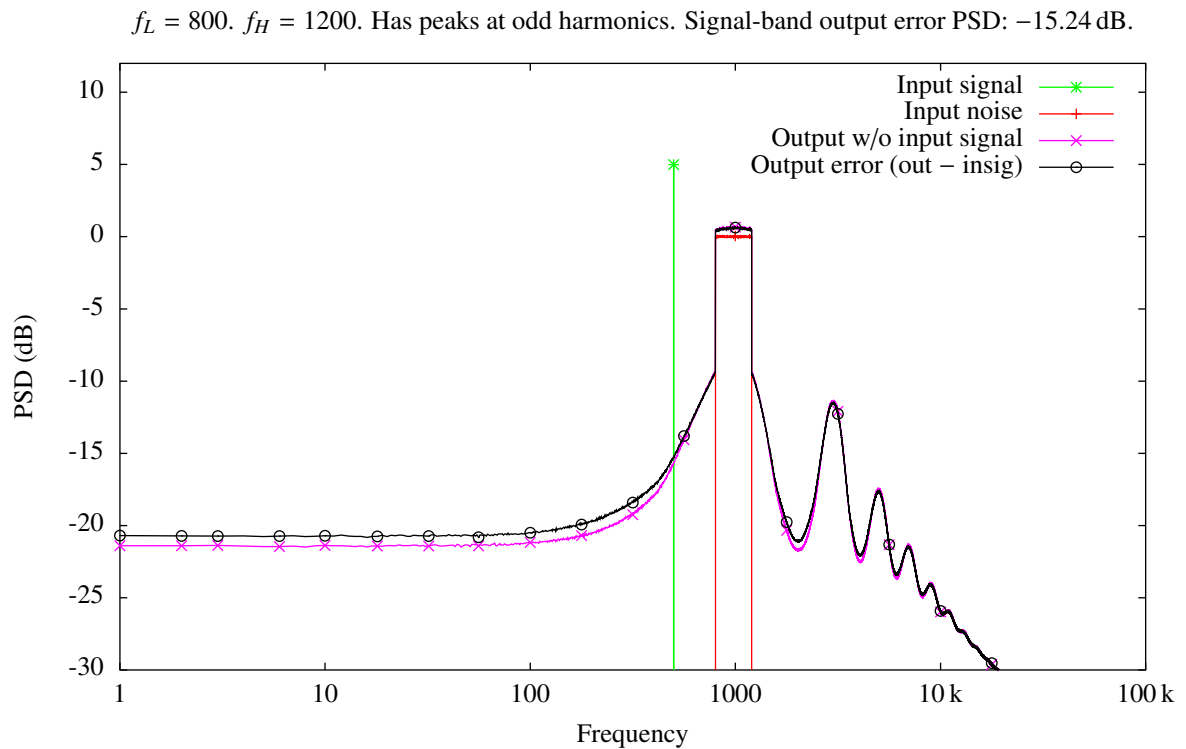


Figure 4.36: Spectrum of continuous-time 1-bit quantization with brick-wall-band-pass-filtered AWGN and signal outside the noise band.

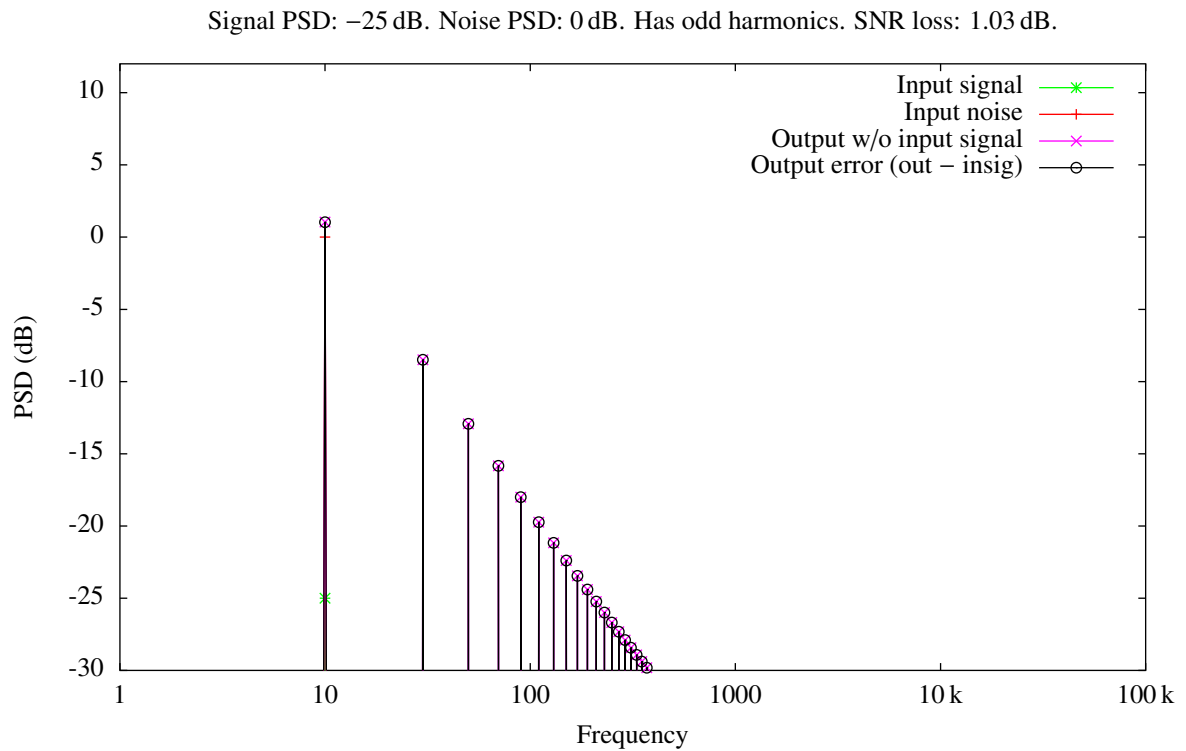


Figure 4.37: Spectrum of continuous-time 1-bit quantization with single-frequency-filtered AWGN.

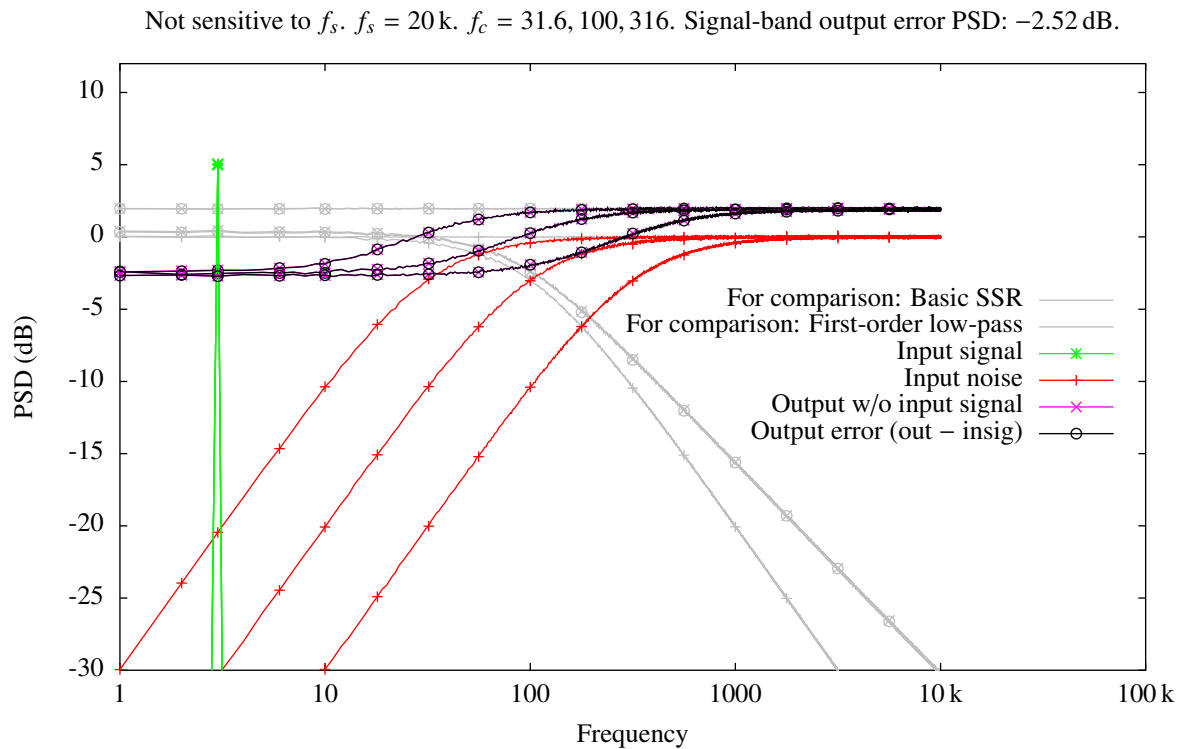


Figure 4.38: Spectrum of sampled 1-bit quantization with first-order-high-pass-filtered AWGN.

Sensitive to f_s . $f_s = 20$ k. $f_c = 2$ k, 5 k, 9 k. Signal-band output error PSD: -3.41 dB, -6.33 dB, -18.01 dB.

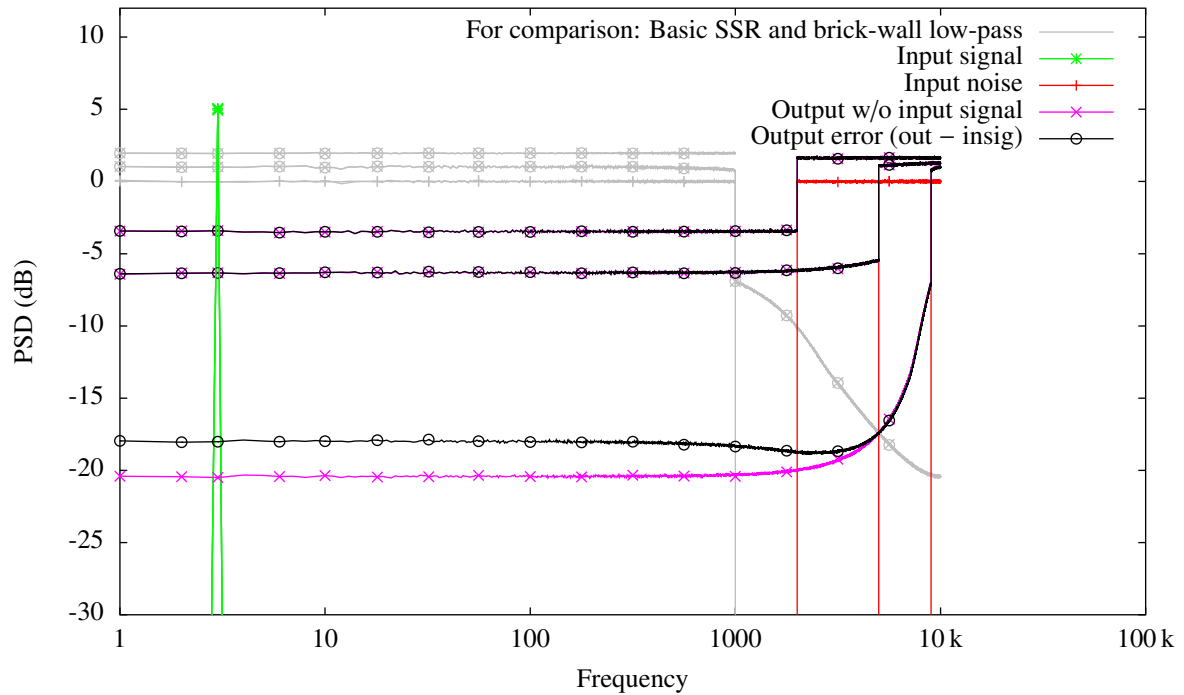


Figure 4.39: Spectrum of sampled 1-bit quantization with brick-wall-high-pass-filtered AWGN.

Sensitive to f_s . $f_s = 20$ k. $f_c = 9$ k. Signal-band output error PSD: -18.01 dB.

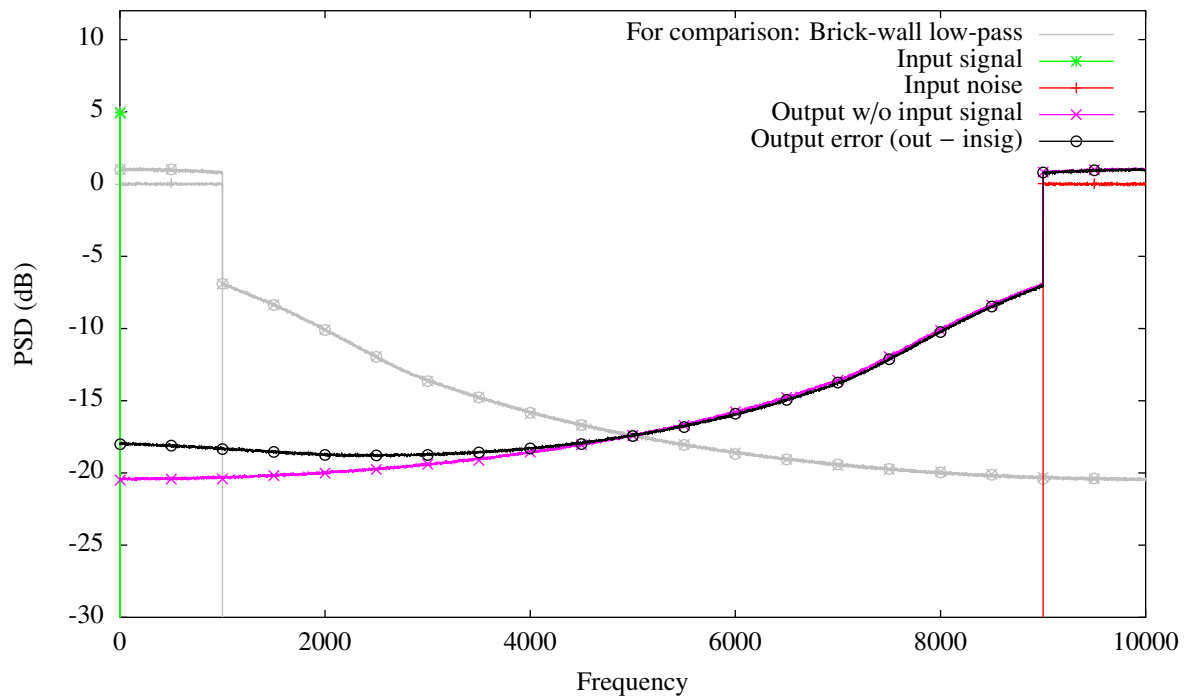


Figure 4.40: Spectrum of sampled 1-bit quantization with brick-wall-high-pass-filtered AWGN, plotted on a linear frequency axis.

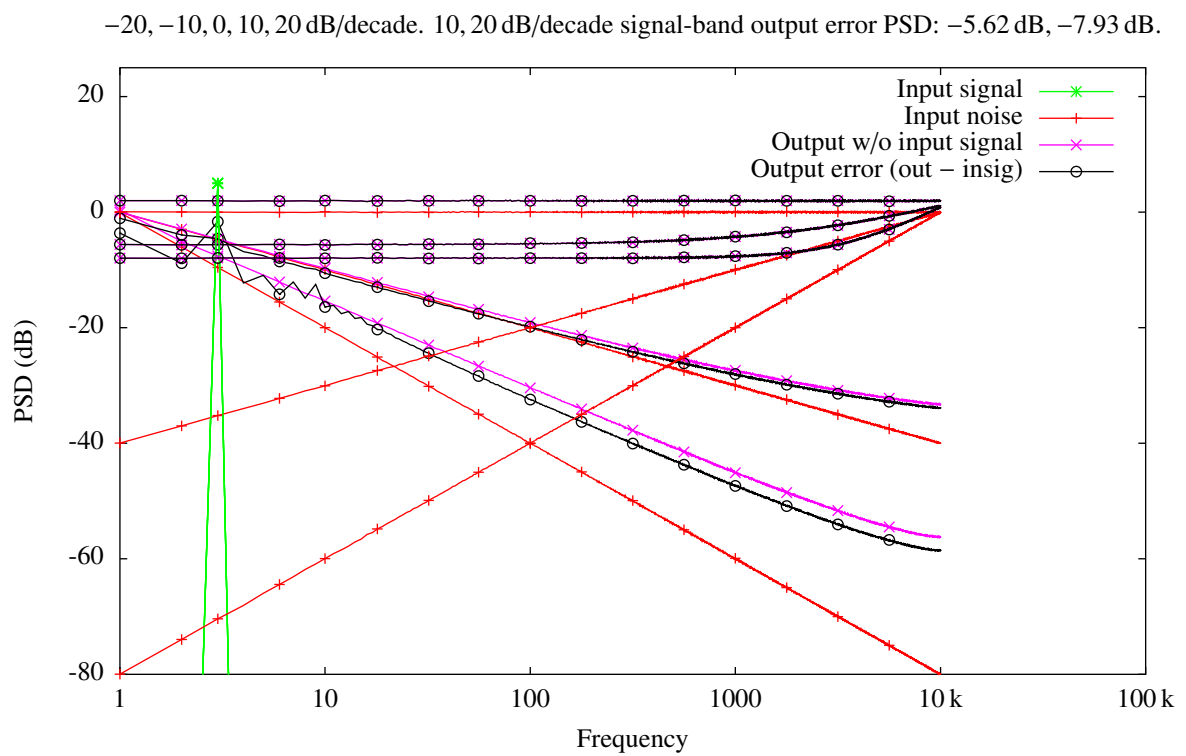


Figure 4.41: Spectrum of sampled 1-bit quantization with power-law-filtered AWGN.

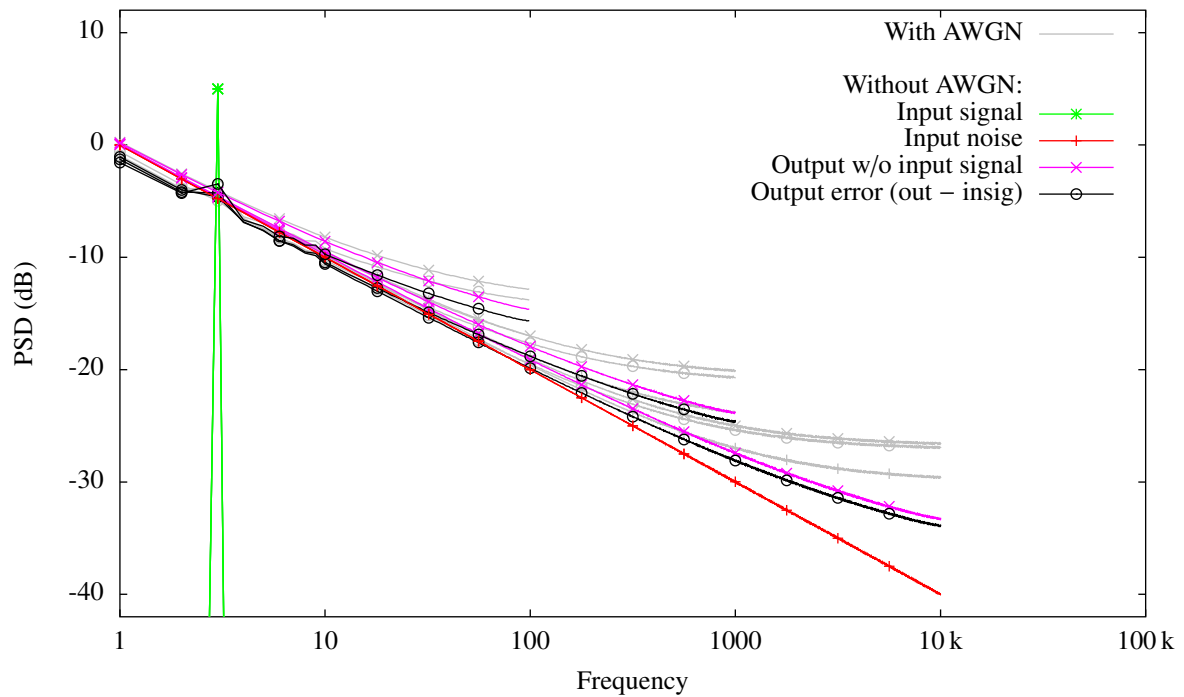


Figure 4.42: Spectrum of sampled 1-bit quantization with flicker noise, i.e. -10 dB/decade power-law-filtered AWGN. Addition of some unfiltered AWGN is also shown.

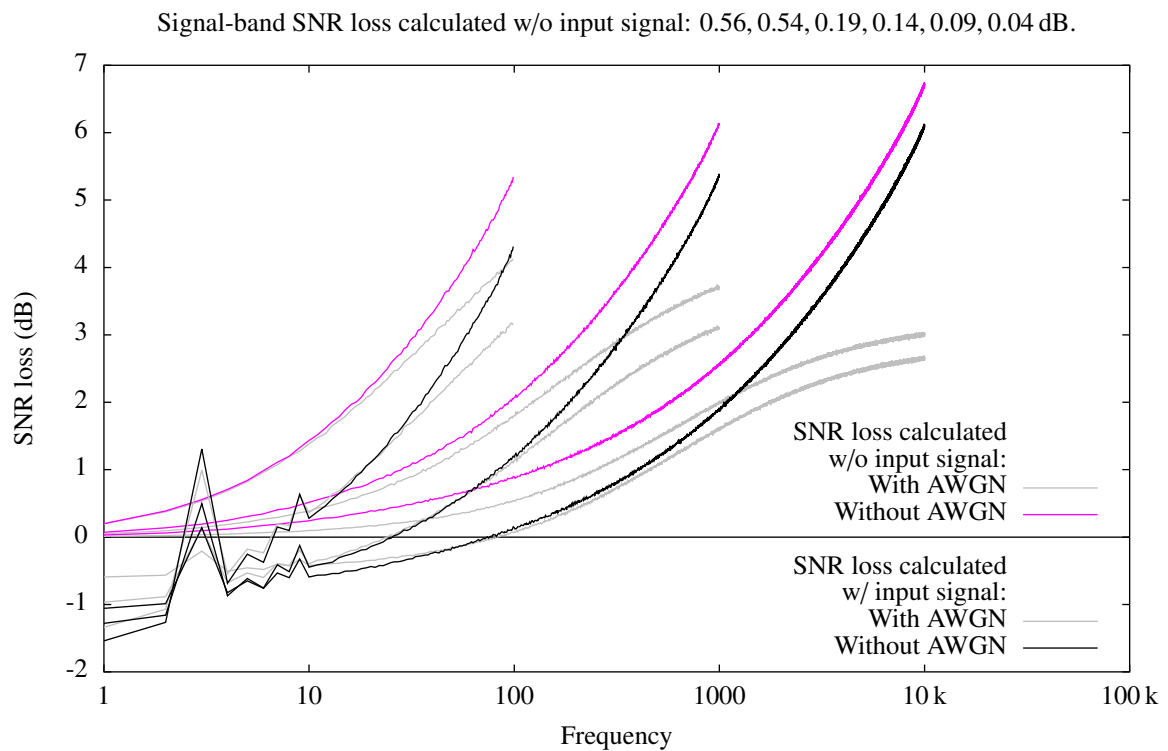


Figure 4.43: SNR loss in the figure above.

Figure 4.44: Triangle (sampled).

Triangle wave with/without white Gaussian noise with std. dev. corresponding to the step size in the triangle (dithering rule-of-thumb from [Widr 96]). This extra noise makes the edges of the PWM pulses a bit fuzzy. Here we see that adding noise can sometimes improve the SNR for the triangle wave. Sampled system. Triangle wave should give good SNR. Noise is placed at high frequencies.

Figure 4.45: Triangle (continuous-time).

Approximation to continuous-time system. The higher the simulation sampling rate is, the better the output SNR is. So fully continuous-time triangle wave should give very good performance.

Figure 4.46: No noise.

Outlevel = $1 \cdot \sigma_{\text{signal}}$ instead of $\sqrt{2/\pi} \cdot \sigma_{\text{signal}}$ for continuous-time system was found by trial and error.

The Signal-to-Quantization-Noise Ratio (SQNR) term could be appropriate here, since the output error only consists of quantization noise. Note that the term is less suited for e.g. a basic SSR system since the error on the output is due to both the input noise and the quantization process, and since the noise power added by the quantization is dependent on the input noise.

Figure 4.47: Sigma-delta.**Figure 4.48, 4.49, 4.50, and 4.51: Noise sweep.**

These are "dumb" systems. They should maybe have changed the quantizer output level factor $\sqrt{\pi/2}$, but instead they keep it the same. This is a feasible scenario if the noise characteristics are known, but the signal characteristics are not. Performance may as a consequence not be optimal at all SNR levels.

The lowered SNR loss is discussed in section 4.11.

Section 4.9.9 discusses why the SNR value can sometimes be higher at the output of a system than at the input.

Figure 4.52, 4.53, and 4.54: Basic SSR with sweep of noise and sampling rate.

Note that the results might be different if other parameters change, such as the input signal bandwidth and PSD.

10x sample rate (actually 3.16x in the plots), -10dB noise, same signal BW: same noise PDF (i.e. same distortion curve) but more integration.

Uniform input noise and uniform input signal which is not stronger than the input noise, or just amplitude-limited input signal, would not exhibit distortion, as shown in figure 4.19, so since the total output power is constant, the output error power and SNR loss would be lower for an input signal with high amplitude than with low amplitude regardless of integration factor. The output error PSD would not hit a distortion floor at high integration factors like it does for Gaussian input noise in figure 4.53.

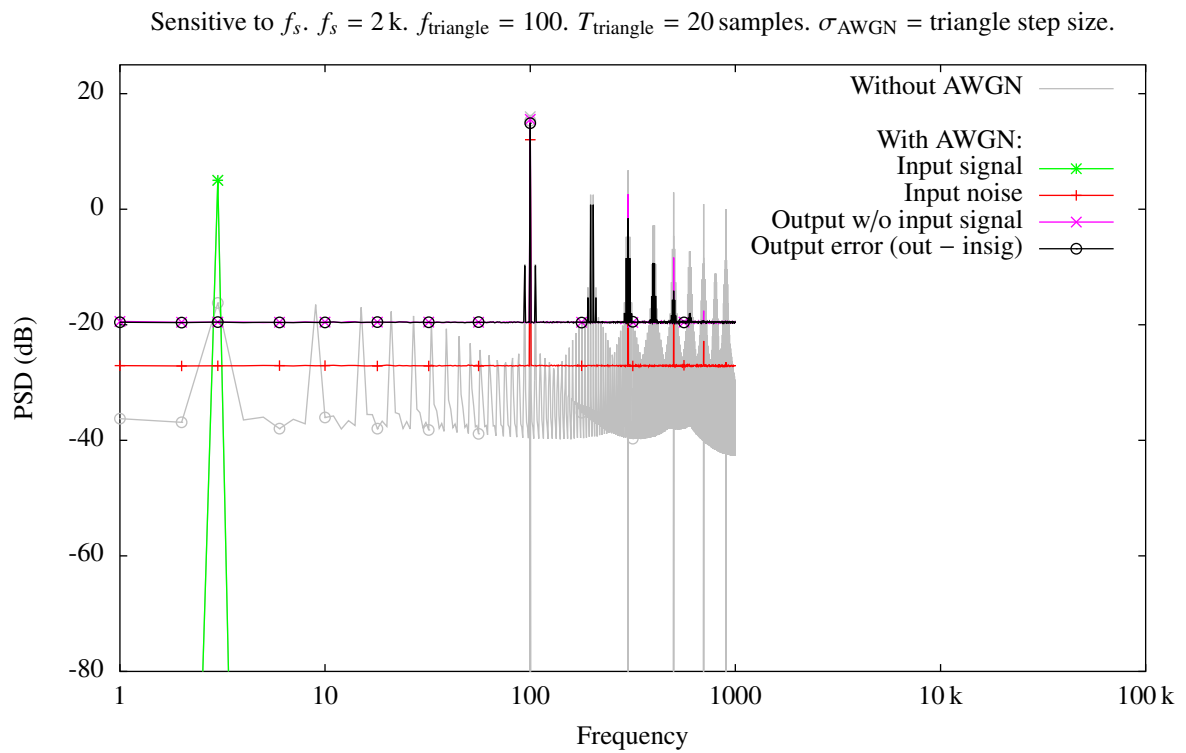


Figure 4.44: Spectrum of sampled 1-bit quantization with a triangle wave as “noise”.

Approximation to continuous-time. Sensitive to simulation f_s . $f_{\text{triangle}} = 100$. $T_{\text{triangle}} = 20, 200, 2000$ samples.

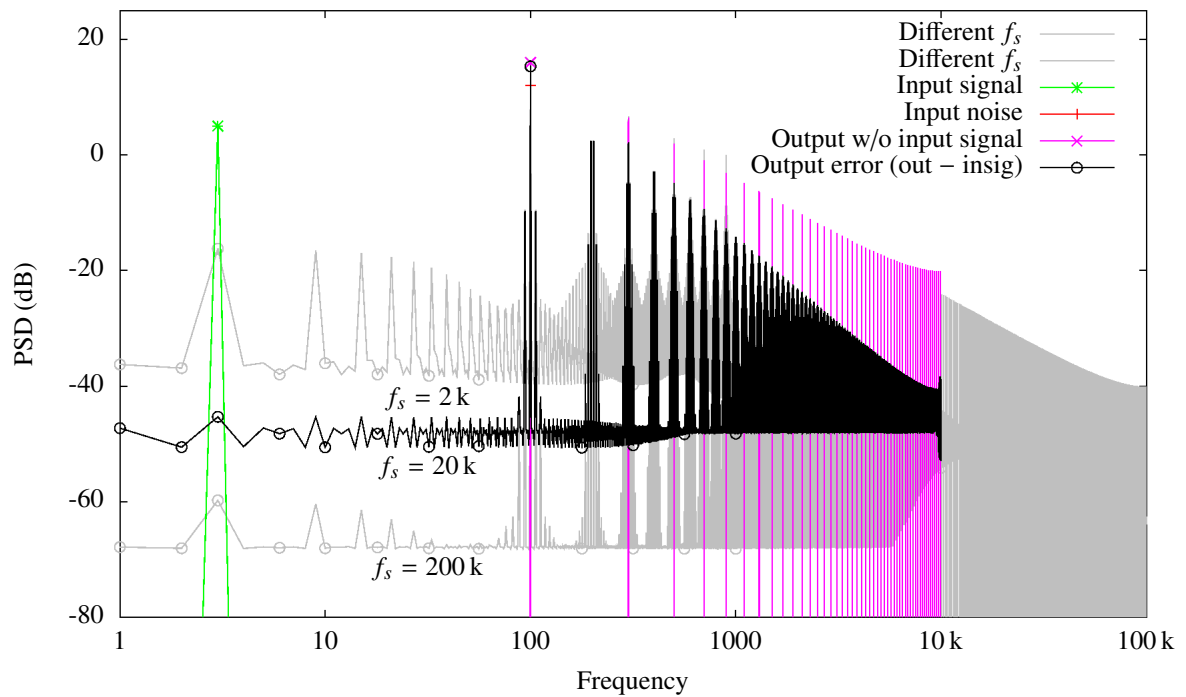


Figure 4.45: Spectrum of continuous-time 1-bit quantization with a triangle wave as “noise”.

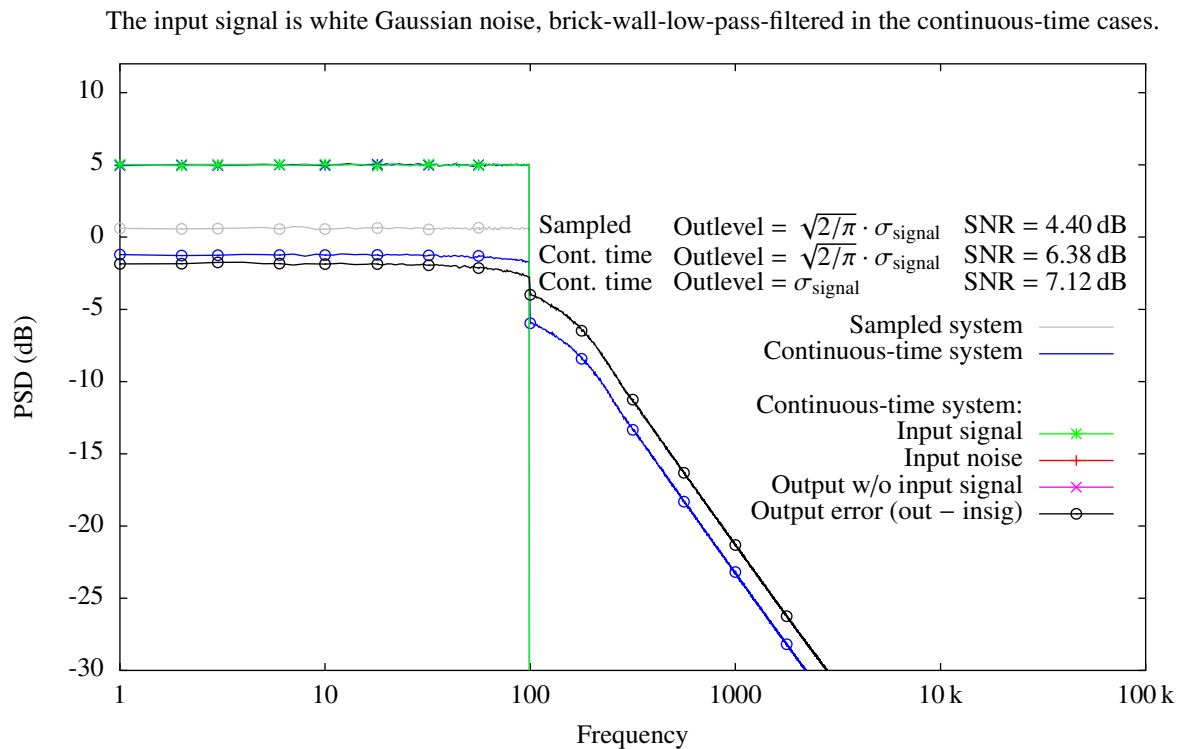


Figure 4.46: Spectrum of sampled and continuous-time 1-bit quantization without any input noise.

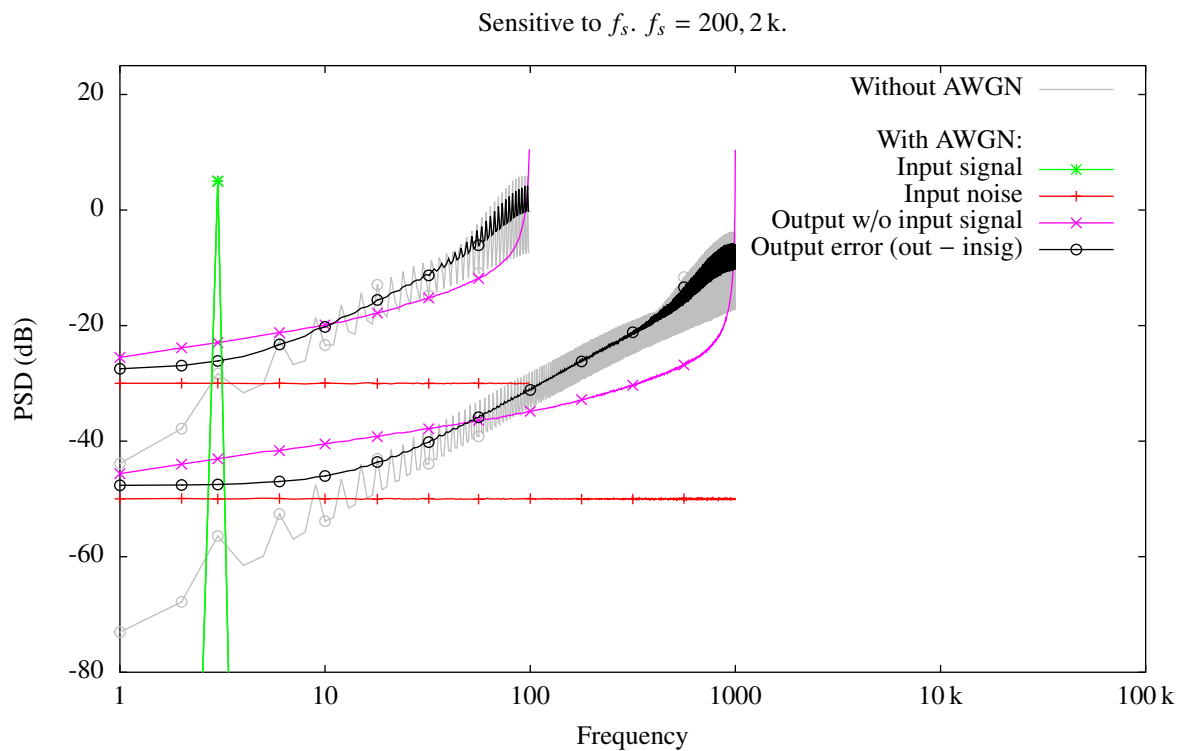


Figure 4.47: Spectrum of sampled 1-bit quantization using a first-order sigma-delta modulator.

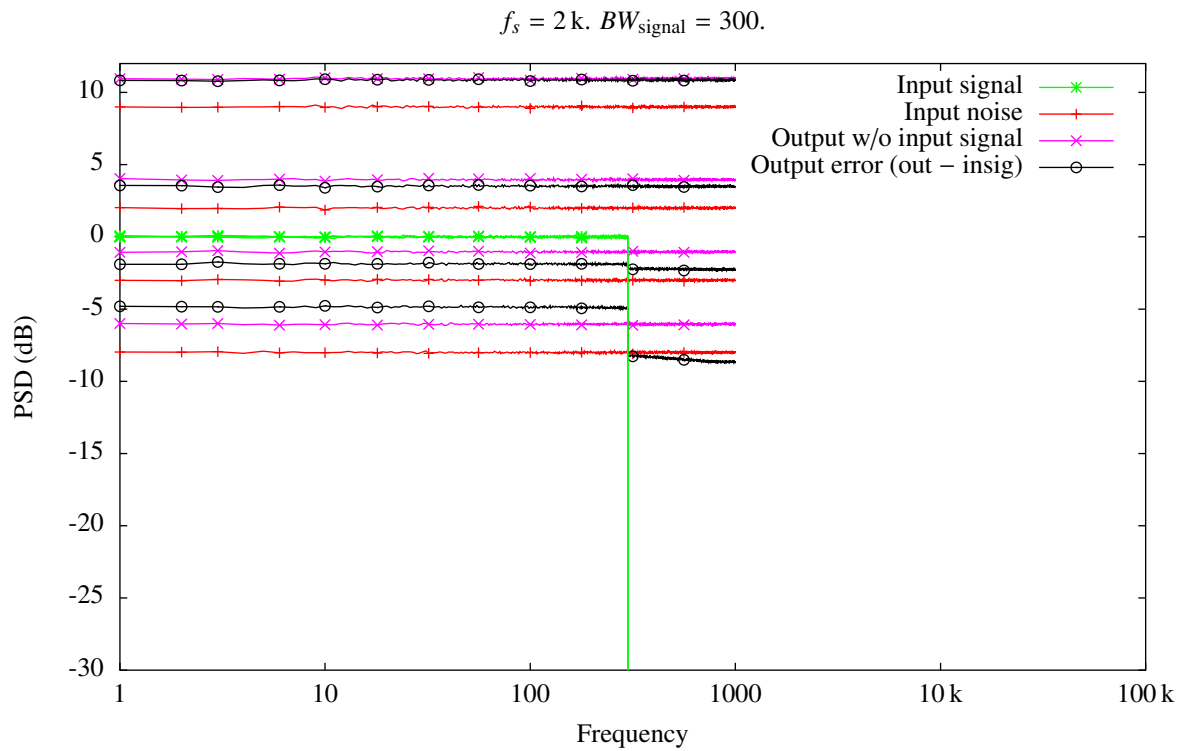


Figure 4.48: Spectrum of sampled 1-bit quantization with AWGN and wideband signal, sweeping the input noise strength.

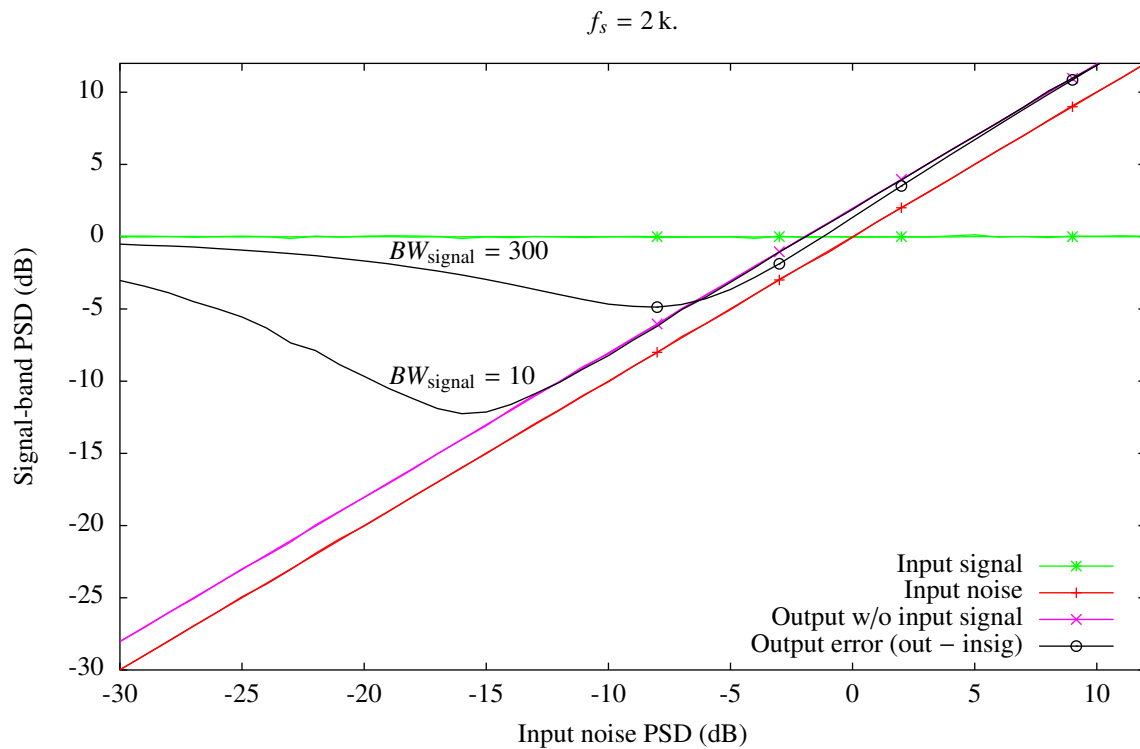


Figure 4.49: Output error strength as a function of input noise strength. The marked points in this figure are plotted in detail in the figure above.

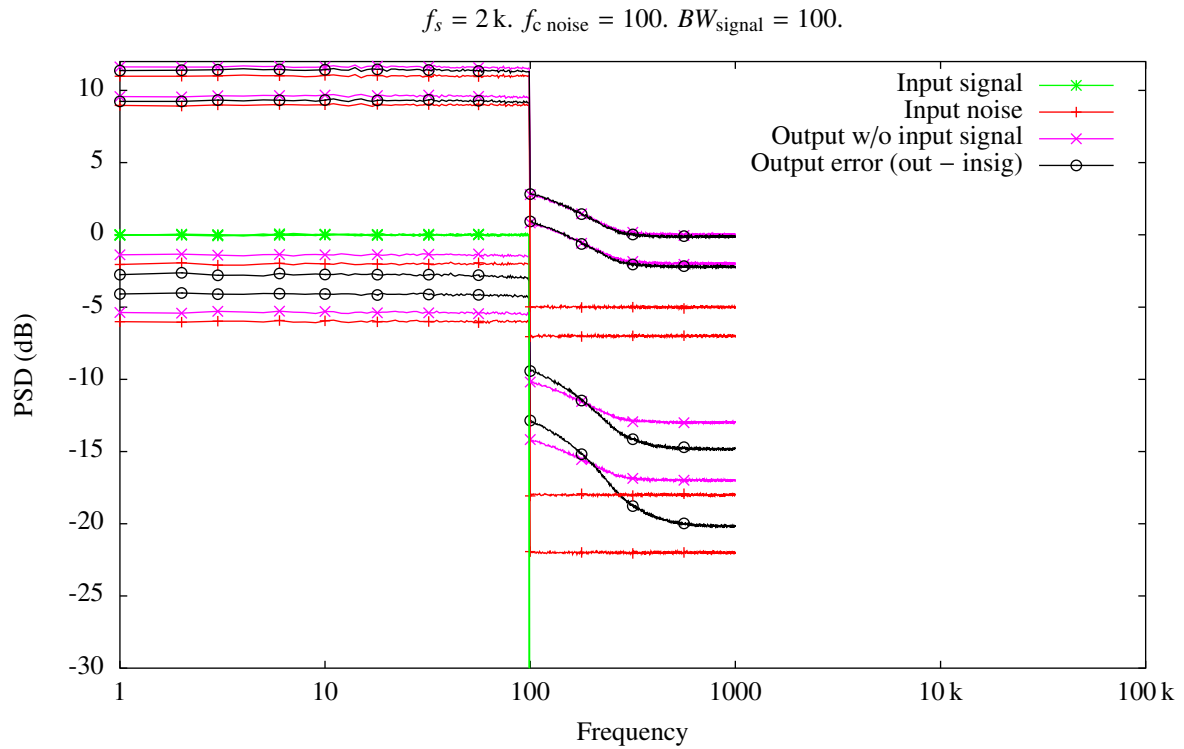


Figure 4.50: Spectrum of sampled 1-bit quantization with high-shelf-filtered AWGN and wide-band signal, sweeping the input noise strength.

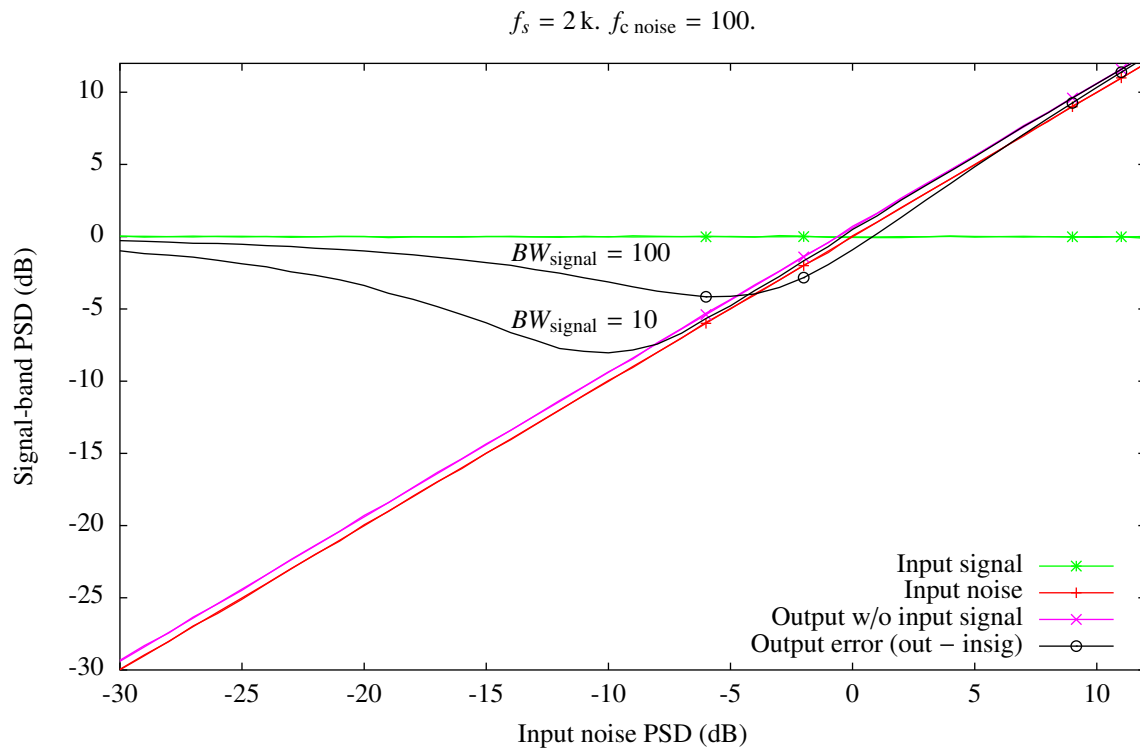


Figure 4.51: Output error strength as a function of input noise strength. The marked points in this figure are plotted in detail in the figure above.

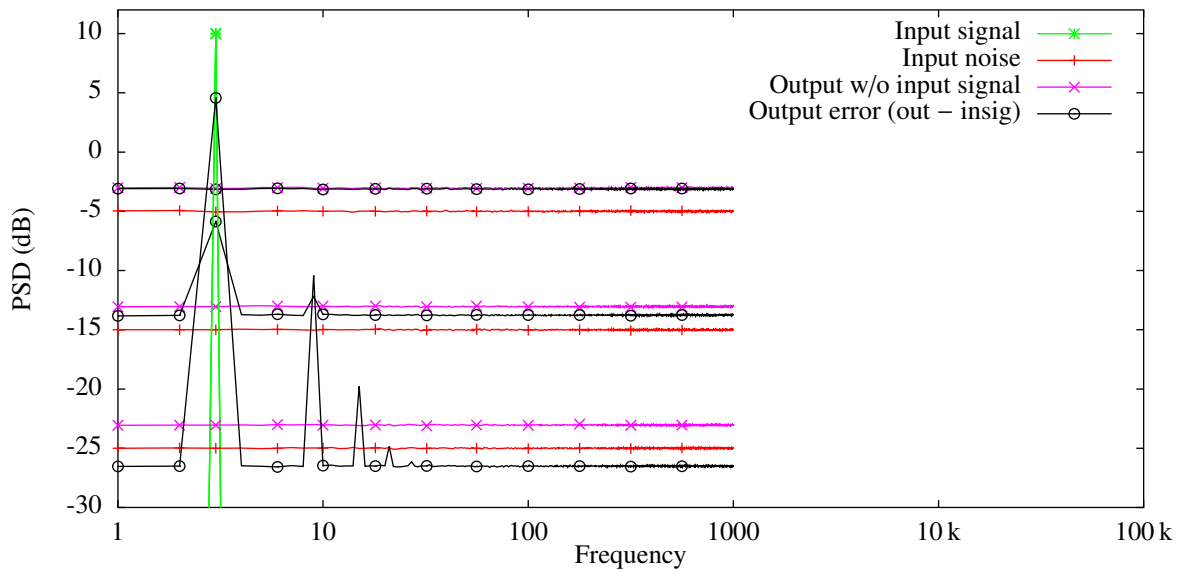


Figure 4.52: Spectrum of sampled 1-bit quantization with AWGN, sweeping noise strength.

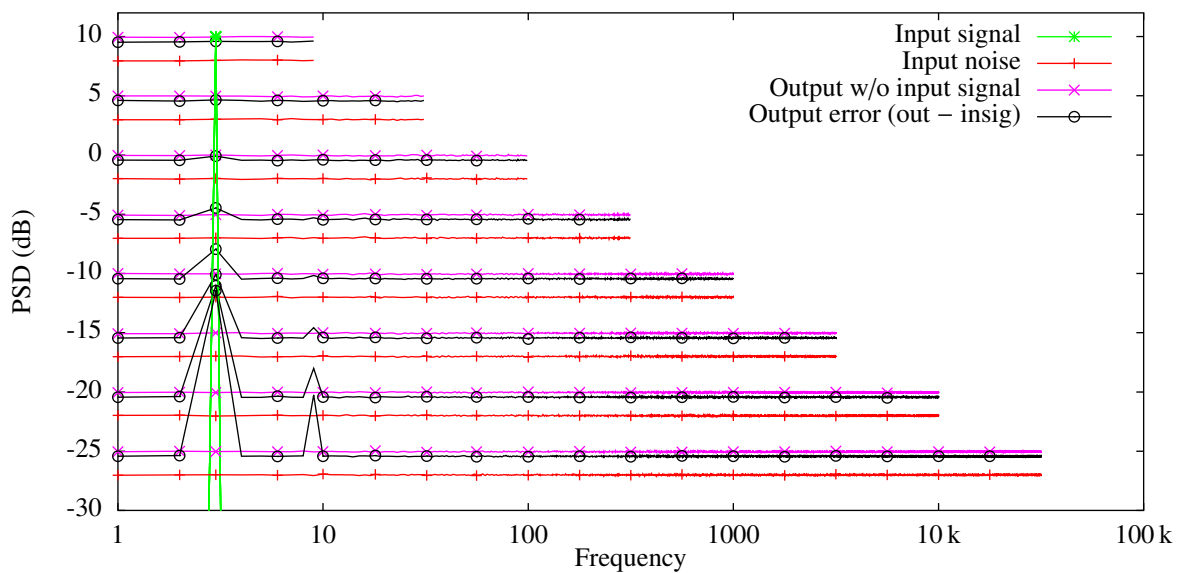


Figure 4.53: Spectrum of sampled 1-bit quantization with AWGN, sweeping noise strength and sample rate, keeping the standard deviation of the noise constant.

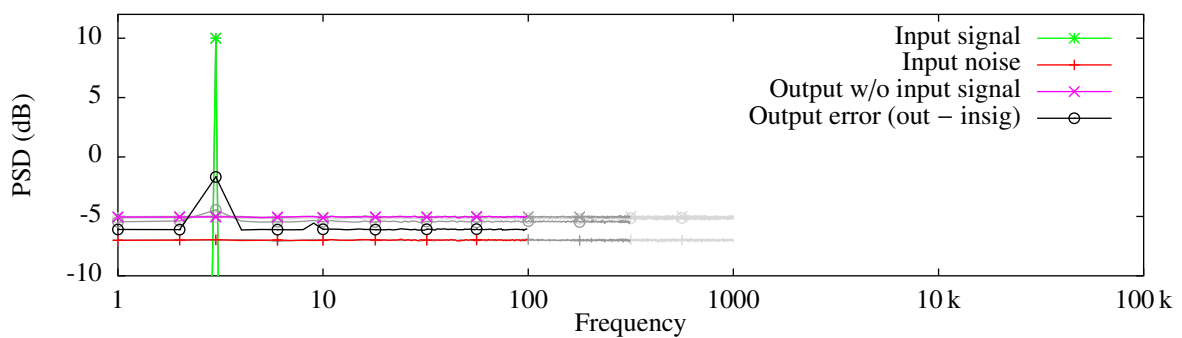


Figure 4.54: Spectrum of sampled 1-bit quantization with AWGN, sweeping sample rate.

4.15 Bitstream Processing

Signal processing can be performed on PDM bitstreams, such as those from SSR and SDM, directly, i.e. without the use of decimation. For example, multiplication of two bitstreams can be done with a simple XNOR gate. Figure 3.3 shows several different signal conversions and signal processing operations that can be performed on bitstreams. Bitstream signal processing has been investigated in literature, some of which we will list below.

It is possible to do bitstream filtering [Case 93 ⑤] and other bitstream processing [East 97 ⑤]. It is also possible to do gain control, mixing (summing), filtering, and equalization [East 97 ⑤]. Adding can be done in a simple way [Fuji 99 ⑤].

Bitstream processing finds applications in 1-bit UWB receivers [Niel 07 ⑤] and PCA detectors for sonar [Kane 66 ⑤].

4.16 Example Applications

1-bit signal processing saves on system cost and power consumption [Zeng 13 ☹]. This is what we have shown in chapter 3 and this chapter.

Example applications for sampled 1-bit signal processing:

- SDM ADCs/DACs
- Direct-Sequence Spread Spectrum (DSSS)/CDMA: Communication over noisy 1-bit or low-SNR channels
- M-sequence radar is basically an SSR system
- Some GPS receivers use 1-bit quantizers [Zeng 13 ☹, Braa 99 ☹]
- 1-bit IR-UWB receiver [Niel 07 ☹]
- Communication receivers with 1-bit processing [Beau 86 ☹]
- 1-bit processing for sonar with Polarity Coincidence Array (PCA) [Kane 66 ☹]

Photographic film consists of many small silver-halide crystals which either receive enough photons to become exposed or remain unexposed, i.e. they have a binary response to light exposure. In other words, photographic film stores the image using oversampled 1-bit values [Wiki 14 ☹].

In an oversampled binary image sensor, smaller and more numerous pixels are used instead of normal-sized pixels. The pixel values are 1-bit quantized, with a threshold level of only a few photons, instead of the usual multi-bit quantization. This makes the sensor resemble photographic film. The 1-bit pixel values are combined in order to find the multi-bit value, using e.g. Maximum-Likelihood Estimation (MLE). The result is a logarithmic-like response function, which makes the sensor suitable for High-Dynamic-Range (HDR) imaging. The oversampling can be spatial or temporal, and Single-Photon Avalanche Diodes (SPADs) can be used as pixels [Sbai 09 ☹] [Wiki 14 ☹].

We might even be able to identify 1-bit signal processing in evolution due to the 1-bit nature of life/death. Assume an organism has an overall genetic fitness $f = f_x + f_{\text{other}}$, where f_x is the genetic fitness of a given feature, e.g. neck length, and f_{other} are the sum of all other features. Also assume that the organism experiences random events in life, r , that influence its chance of survival. Assume survival of the organism, and thus also procreation, is determined by $f + r > 0$, i.e. $f_x + f_{\text{other}} + r > 0$. We then see that we have a small signal plus a lot of noise, which is then 1-bit quantized. The next thing we need is integration. The survival status of e.g. a billion organisms should provide plenty of that. By multiplying the survival data with a $-1/1$ array describing which gene variant each organism has, we should get a value out which says something about the fitness of the two gene variations. This is basically like CDMA demodulation, where many different channels can be picked up from a single bitstream, as in a 1-bit GPS receiver.

Chapter 5

Continuous-Time 1-Bit Signal Processing

Contents

5.1	Introduction	138
5.1.1	Example Applications	138
5.2	Zierhofer Modulation	140
5.3	CTBV-Based Logic	141

5.1 Introduction

1-bit signals can also be continuous-time. We have called this Continuous-Time Binary-Value (CTBV) [Hjor 06 ☞] [Hjor 09 ☞]. This kind of signals has also been investigated elsewhere in literature [Sche 08 ☞]. Zierhofer modulation and RALe Logic Architecture (RALA) are two examples of interesting preexisting uses of CTBV signals, and these two concepts are discussed in section 5.2 and 5.3 below. Figure 5.1 illustrates how CTBV signals work. We will not perform an extensive theoretical analysis of CTBV signals in this thesis. Instead, we will present many different practical uses for such signals in CMOS circuits. These solutions are shown in the papers in appendix A.

5.1.1 Example Applications

Example applications of continuous-time 1-bit signals:

- PWM control of LED lights
- PWM control of RC servos
- Class-D amplifier with PWM
- Half-tone printing method
- Nerve signals

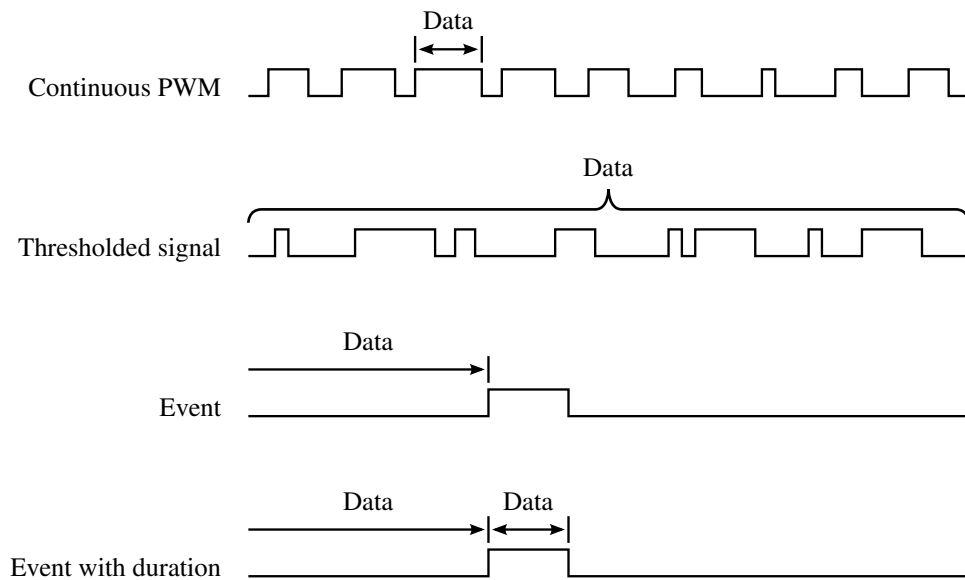


Figure 5.1: Some various types of Continuous-Time Binary-Value (CTBV) signals. Note how the data values are conveyed by continuous-time transition events and thus have high precision even though CTBV signals are very simple and coarsely quantized. For the thresholded signal, it is not the values but rather the time axis of the data which is continuous, though.

5.2 Zierhofer Modulation

There exists a modulation created by Zierhofer [Zier 12 ©] which improves upon basic Pulse-Frequency Modulation (PFM). It requires fewer impulses than PFM and is completely lossless, but on the other hand it is more complicated and does not convey the DC level of the signal.

The analysis of the modulation method considers a periodic signal with B harmonics and no DC component, i.e. a signal with frequency components $\{1, 2, 3, \dots, B\}$. The signal must also obey rules concerning its amplitude. The conversion process is somewhat complicated, but eventually yields $2B + 1$ unity-weight Dirac impulses with specific timings. These timings are real-valued, or in other words, the modulation has continuous-time output. This is basically a PDM modulation, so the original signal can be recovered using an ideal low-pass filter. The modulation is completely lossless.

For comparison, the same signal, but including the DC component, can be represented in the frequency domain as a DC component and B phasors. This corresponds to $2B + 1$ real values or degrees of freedom in terms of data content. The signal can also be sampled according to the Nyquist-Shannon sampling theorem with a sampling rate $f_s > 2B$, e.g. $f_s = 2B + 1$, which means that also this representation would require $2B + 1$ real values of data. With the exception of the missing DC component, the modulation method by Zierhofer is therefore a one-to-one mapping between the original signal and its representation since it has the same degrees of freedom as the frequency domain or sampled representations. In other words, this modulation is lossless and it does not add any unnecessary information content in the conversion process except one real value. So in terms of number of impulses required, this representation uses just one more than as low as it could possibly go, i.e. the $2B$ real values of the original signal are represented using $2B + 1$ impulses with real-valued timings.

The extra degree of freedom is probably the position of the first impulse. Moving it would shift the other impulses too, and the signal would thus still be represented correctly, but with a different sequence of impulses.

Note that, unlike frequency-domain or sampled representations, the real values in the impulse timings are bounded since the signal is periodic. This is probably the underlying reason why the original signal must follow amplitude constraints.

The modulation method can also operate on non-periodic signals, but in order for this running conversion process to be causal, the converter will have to delay the signal by a certain amount. In other words, the position of an impulse depends not only on the signal up to that point, but also on what the signal will do next. Therefore, the converter needs to be able to see a little ahead in time.

As is well known, a band-limited signal can be represented without loss using equidistant weighted Dirac impulses. It is interesting to note that the modulation described here does the exact opposite, in that it uses unity-weight Dirac impulses with specific positioning, also here to represent the signal without loss. Both representations yield the original signal when passed through an ideal low-pass filter.

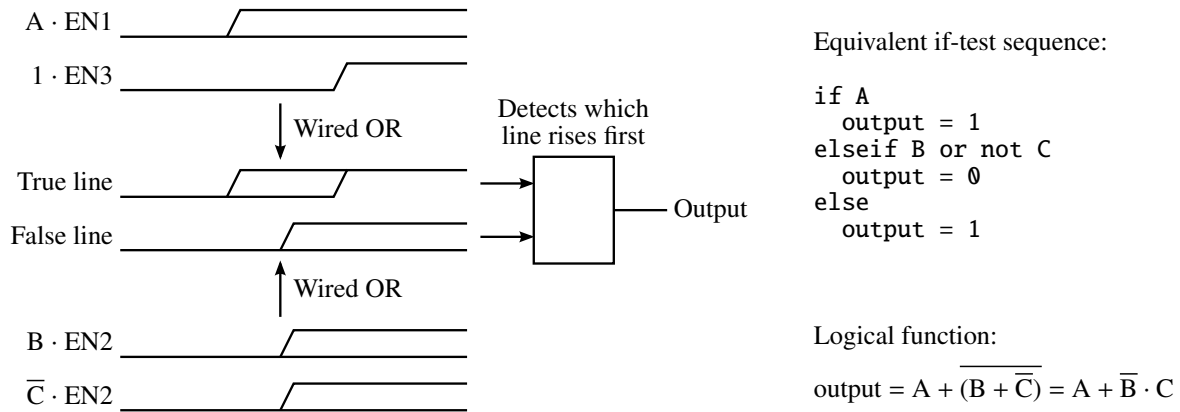


Figure 5.2: Exemplified concept of RALe Logic Architecture (RALA), a logic family which uses CTBV to perform the calculations.

5.3 CTBV-Based Logic

A novel logic family named RALe Logic Architecture (RALA) has been proposed by Lee and Yoo [Lee 02 ©]. This logic family uses CTBV to perform the calculations, i.e. it uses signal timing as a computational element. It is only the internals of the logic circuit which uses CTBV; the inputs and the outputs are normal logic values. The logic family is claimed to have several advantages over other logic families.

Figure 5.2 exemplifies the concept behind RALA. The Boolean inputs A, B, and C are ANDed with enable signals ENx. One after another, these enable signals transition from 0 to 1, thus letting through the inputs in a given sequence. The signals are then connected together with two wired-OR circuits. One wired-OR output is called the true line and the other is called the false line. A circuit then detects which line rises first. If the true line rises first, the output of the RALe circuit is true, and if the false line rises first, the output is false. The operation of the circuit can be seen as an if/elseif/else sequence, as shown in the figure. Such an if-test sequence terminates on the first match, just like the RALe output is determined solely by the first transition of the true/false lines. The end result is that the output is a logical function of the inputs. Arbitrary logical operations can be implemented, but sometimes a cascade of transition detectors is required.

Chapter 6

Circuit Solutions

Contents

6.1	Introduction	144
-----	------------------------	-----

6.1 Introduction

The published papers in appendix A cover a wide range of technologies. The common thread is that most of them use 1-bit signal processing and UWB, and that most of them are also useful in WSN applications. WSN, UWB, 1-bit, and CMOS work well together because they shared properties such as simplicity, low price, low SNR, and high bandwidth. The technologies that are implemented in the papers, are summarized below:

- Radar
 - Swept-threshold sampling
- Localization
 - Clock burst generator
 - CTBV symbol detector
- Communication
- Beamforming
 - Delay calibration
- Bitstream processing
- General circuits
 - Quantizer
 - Pulse generator
 - N-path band-pass filter
 - Power harvesting
 - SPI

Figure 6.1 summarizes the subjects covered in the papers and their relationships in a graphical manner.

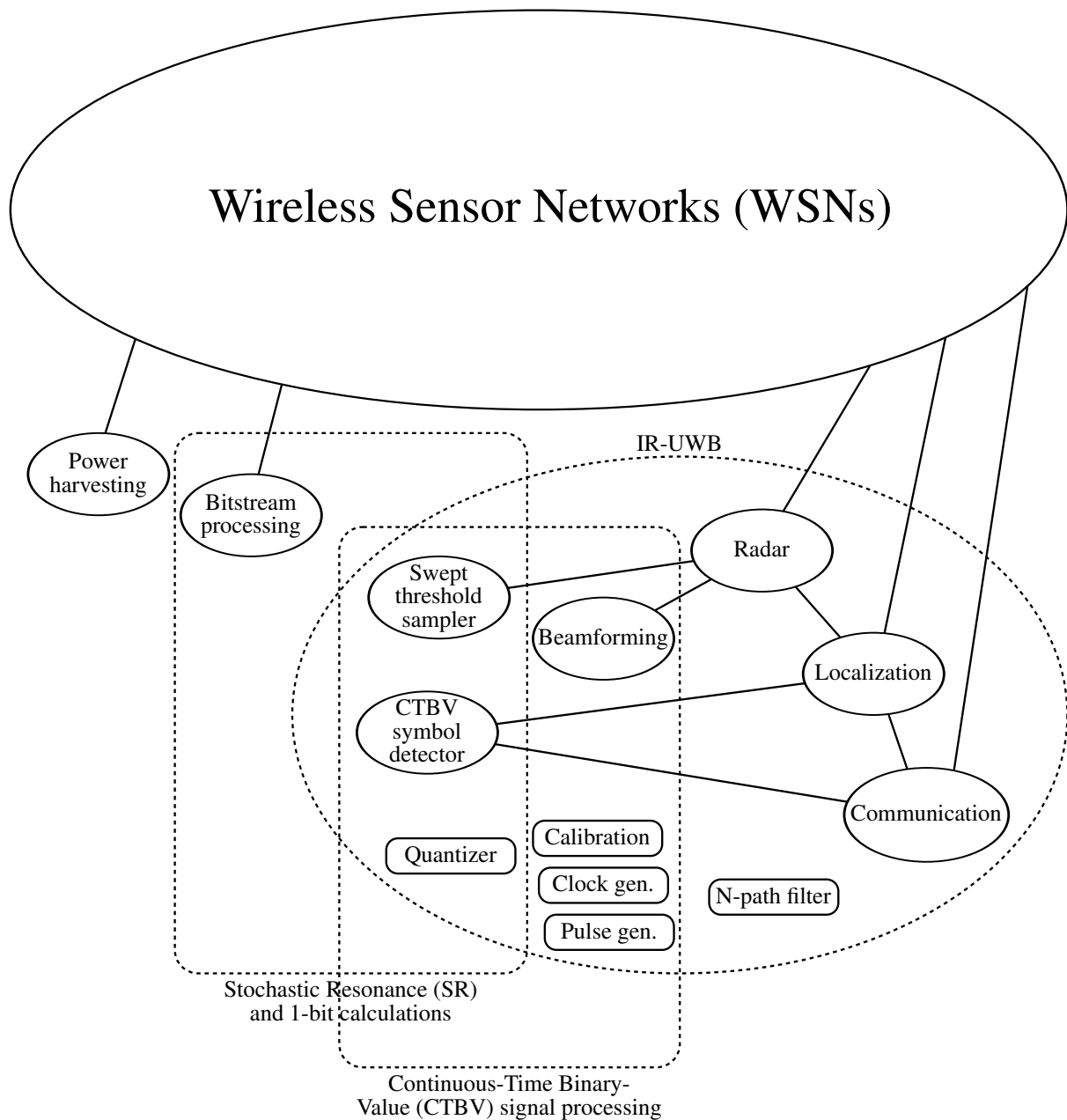


Figure 6.1: Overview of subjects covered in the papers.

Chapter 7

Conclusions

Contents

7.1	Conclusions	148
7.2	Main Contributions of This Work	148
7.2.1	General Overview of 1-Bit Signal Processing	148
7.2.2	Classification of 1-Bit Concepts	148
7.2.3	Energy Consumption per Sample	148
7.2.4	Power Comparison of Modulations/Encodings	149
7.2.5	Forced Oversampling in Analog Systems	149
7.2.6	1-Bit Signal Processing As a Fundamental Principle	149
7.2.7	SNR Metrics	149
7.2.8	Extra Strong SR Peak	149
7.2.9	Large-Signal Analysis	149
7.2.10	SNR As a Function of Input Noise	149
7.2.11	Why SSR with Colored Noise Works	150
7.2.12	White Noise from First-Order Anti-Aliasing Filter	150
7.2.13	Spectrum Simulations	150
7.2.14	Practical SSR System	150
7.2.15	Flicker Noise Gives 0 dB SNR Loss	150
7.2.16	Continuous-Time SSR Without Noise	150
7.2.17	SSR in Evolution	151
7.2.18	CTBV Symbol Detector	151
7.2.19	High-Precision Timing Calibration	151
7.3	Future Work	151

7.1 Conclusions

We have shown that 1-bit signal processing shows great promise as a system architecture, at least for some types of applications.

In chapter 3, we have tried to get to the deep of what 1-bit signal processing generally is and what its properties are. We have explored different usage classes and given an overview of ways to represent analog signals. We have also tried to classify different concepts related to 1-bit signal processing in order to facilitate insight. Furthermore, we analyzed the power consumption of 1-bit systems and compared it to other classic systems.

In chapter 4, we explored the SSR system, which has already been investigated extensively in literature. We explained when adding noise and nonlinearity to a system might actually improve it. We thoroughly explored three different SNR metrics, and used them throughout the chapter. We performed a small-signal and a large-signal analysis of the SSR system, and plotted SNR as a function of noise. Finally, we explored colored noise and performed many spectrum simulations with different colors of noise and also with other variations.

1-bit signal processing is an interesting option for the CMOS/UWB/WSN combination. This has been demonstrated in silicon, as shown in a range of published papers listed in appendix A.

When evaluating potential architectures such as analog, Nyquist-rate multi-bit, and 1-bit for an application, properties such as signal quality, processing features, implementation cost, and power consumption will be deciding factors for which architecture to choose.

One of the most interesting results in this thesis concerns flicker noise, and is explained in section 7.2.15 below.

7.2 Main Contributions of This Work

Here is a list of interesting results from this thesis that seem to be novel as far as we can tell.

7.2.1 General Overview of 1-Bit Signal Processing

The general overview of 1-bit signal processing that chapter 3 is, seems to not have been done before.

7.2.2 Classification of 1-Bit Concepts

We have classified 1-bit concepts as underlying data carriers, basic modulation types, and modulators. This is discussed in section 3.2 and illustrated in figure 3.1 and 3.3.

7.2.3 Energy Consumption per Sample

We have found that the energy consumption when conveying one sample through a system is given by $E = DR \cdot (kT/2) \cdot P_r$, where DR is dynamic range and P_r is the relative power consumption factor for a given amplifier. This is discussed in section 3.4.1 and 3.4.2.

7.2.4 Power Comparison of Modulations/Encodings

We have analyzed the power consumption of analog, multi-bit, and 1-bit modulations/encodings based on the dynamic range and number of samples they require from the layer below. We found that for some cases with low/moderate SNR, the simple 1-bit SSR systems can compete with analog and classical digital systems. This is discussed in section 3.4.3.

7.2.5 Forced Oversampling in Analog Systems

We introduced the term “forced oversampling” to describe the fact that analog system are always on. This “oversampling” could cause more signal integration and thus a higher power consumption than desired, and might be difficult to get rid of. This is discussed in section 3.4.3.4.

7.2.6 1-Bit Signal Processing As a Fundamental Principle

We have made a connection between 1-bit signal processing and photons, electrons, and quantum physics. This suggests that 1-bit signal processing might be a fundamental principle in nature. This is discussed in section 3.6.

7.2.7 SNR Metrics

We have introduced and analyzed three different SNR metrics. These have probably been used before, but our extensive investigation of them might have provided some new insights. This is discussed in section 4.9.

7.2.8 Extra Strong SR Peak

We have investigated an effect where the output noise in an SSR system becomes lower than one might expect. Without knowledge of this phenomenon, it may cause some confusion, as we have also exemplified. It is hard to search for earlier work on this since the effect is quite specific yet fuzzy. It is still worth mentioning here as a potential novel insight, though. This is discussed in section 4.11.1.

7.2.9 Large-Signal Analysis

We have performed a visualized large-signal analysis that decomposes the output error into distortion error and random error. The effect of integration is also shown in a very visual and accessible way. We also show what the small-signal SNR loss is at each value of the input signal. This does not seem to be a very common thing to do, but it gives us some interesting insights about the SSR system. This is discussed in section 4.11.

7.2.10 SNR As a Function of Input Noise

We have plotted the output SNR of an SSR system as a function of input noise strength using our three SNR metrics. We have also plotted decompositions of the output noise into random

noise and distortion error. Finally, we have also plotted quantizer output levels, i.e. decoding, as a function of input noise strength. An interesting feature of the plots is that we clearly see how the SNR is asymptotically 1.96 dB lower than an ideal analog system for high input noise. Many of these plotted features seem to be novel. Some SNR metric B plots have been done before, and so have SNR plots for something close to SNR metric C, as explained in more detail in section 4.12.3. This is discussed in section 4.12.

7.2.11 Why SSR with Colored Noise Works

We have proved why SSR with colored noise will work just as well as with white noise, since this is not immediately obvious. This is discussed in section 4.13.1.

7.2.12 White Noise from First-Order Anti-Aliasing Filter

First-order low-pass filtered noise can be made white again by sampling it a certain way. The reason for this is aliasing effects. This is discussed in section 4.13.2.

7.2.13 Spectrum Simulations

We have performed simulations that show the spectrum of the output noise from systems with continuous-time thresholding, i.e. SSR systems without sampling. We have done many simulations with many parameter variations, but mostly with different noise colors. Although the main results, such as SNR loss, from these simulations have been presented in earlier works, the full spectrum plots themselves seem to be novel. This development might be due to the increasing availability of powerful computing. This is discussed in section 4.14.

7.2.14 Practical SSR System

We have designed a practical SSR system with a third-order filter which gives very low SNR loss. No other anti-aliasing filter than this is required. This is discussed on page 117 in section 4.14.3.

7.2.15 Flicker Noise Gives 0 dB SNR Loss

An SSR system with flicker noise, i.e. noise with a -10 dB per decade roll-off, has close to 0 dB SNR loss. This seems to be a very interesting and novel result, and might be one of the most interesting results in this thesis. This is discussed on page 122 in section 4.14.3.

7.2.16 Continuous-Time SSR Without Noise

It is a known result that when there is no noise in an SSR system, the quantizer output levels should be set to $\sqrt{2/\pi} \cdot \sigma_{\text{signal}}$. We have demonstrated that $1 \cdot \sigma_{\text{signal}}$ is a better value for the continuous-time case. This is discussed on page 128 in section 4.14.3.

7.2.17 SSR in Evolution

We have identified a possible SSR mechanism in the process of evolution itself, not just being used by biology. The reason for this is that genetics and circumstance can be seen as signal and noise, and that life/death can be seen as a 1-bit quantization. This is discussed in section 4.16.

7.2.18 CTBV Symbol Detector

We have created in silicon an IR-UWB symbol detector that is based on Continuous-Time Binary-Value (CTBV), which is a very unusual way of creating a system. This is discussed in the papers in appendix A.2.

7.2.19 High-Precision Timing Calibration

We have created a digital delay element and a calibration circuit that together can yield very precisely controlled time delays. This is discussed in paper 15.

7.3 Future Work

The following are interesting projects which could be attempted in the future.

The energy per degree of freedom for the frequency components seems to be a bit more difficult to work with when the load is capacitive than when it is resistive, as described on page 44 in section 3.4.1.5. This could be interesting to look a bit closer into. There are also several questions about the general $kT/2$ rule which it could be interesting to look into too, as described in section 3.4.1.6.

SNR measurement in oversampled signals with distortion: When the input signal is narrow-band-filtered white Gaussian noise, it is a sinusoid with different amplitude for each simulation iteration, i.e. it does not have a Gaussian distribution during a single simulation run. Look into the consequences of this.

SNR metrics B and/or C for the 1-bit quantization spectrum simulations.

Secondary SNR effects such as intermittent signal fallout due to very-low-frequency components of flicker noise.

It could be interesting to try to implement a symbol detector using sampled SSR, and also to integrate the energy from all of the multipath components.

An M-sequence radar is basically an SSR system, so it could be interesting to try to implement this too.

Appendix A

Papers

Contents

A.1 Radar and CTBV	155
Paper 1 CMOS Impulse Radar	157
Paper 2 Thresholded samplers for UWB impulse radar	165
Paper 3 Impulse Radio technology for Biomedical applications	171
Paper 4 CTBV Integrated Impulse Radio Design for Biomedical Applications	177
Paper 5 CMOS nanoscale impulse radar utilized in 2-dimensional ISAR imaging system	189
A.2 Localization and Communication	197
Paper 6 Power-efficient CTBV symbol detector for UWB applications	199
Paper 7 Continuous-time single-symbol IR-UWB symbol detection	205
Paper 8 Continuous-time high-precision IR-UWB ranging-system in 90 nm CMOS	211
Paper 9 Live demonstration: IR-UWB transceiver with localization based on ToF technique suitable for wireless capsule endoscopy	217
Paper 10 Continuous-time symbol detector for IR-UWB rake receiver in 90 nm CMOS	221
Paper 11 A continuous-time IR-UWB RAKE receiver for coherent symbol detection	227
Paper 12 An IR-UWB Transmitter for Ranging Systems	241
A.3 Beamforming	249
Paper 13 Close range impulse radio beamformers	251
Paper 14 Electromagnetic impulse radio camera	259
Paper 15 High precision calibrated digital delay element	267
Paper 16 An ultra-wideband receiving antenna array	271
Paper 17 A linear IR-UWB MIMO radar array	277
A.4 Bitstream Processing	287
Paper 18 Power efficient cross-correlation beat detection in electrocardiogram analysis using bitstreams	289
Paper 19 Power efficient cross-correlation using bitstreams	295
Paper 20 Power-Efficient Cross-Correlation Beat Detection in Electrocardiogram Analysis Using Bitstreams	301
A.5 Pulse Generator, N-Path Filter, Quantizer, and Power Harvesting	311
Paper 21 A novel 6.5 pJ/pulse impulse radio pulse generator for RFID tags	313
Paper 22 A 5.2 pJ/pulse impulse radio pulse generator in 90 nm CMOS	319
Paper 23 An inductorless 6-Path band-pass filter with tunable center frequency for UWB applications	325
Paper 24 Continuous-time CMOS quantizer for ultra-wideband applications	331
Paper 25 A 70-dB, 3.1-10.6-GHz CMOS amplifier in low-power 90 nm CMOS	337
Paper 26 Power Harvesting Circuits in 90 nm CMOS	343

Introduction

The published papers that are included here are largely about circuits that utilize 1-bit signal processing. Some other Wireless Sensor Network (WSN) technologies such as power harvesting are also included, though. See the list of publications on page 7 for a chronologically sorted list of the papers. Note that the original page numbers in the papers have been overstruck in order to avoid confusion with the page numbers of the thesis.

A.1 Radar and CTBV

Paper 1 CMOS Impulse Radar

H. A. Hjortland, D. T. Wisland, T. S. Lande, C. Limbodal, and K. Meisal. “CMOS Impulse Radar”. In: *Norchip Conference, 2006. 24th*, pp. 75–79, Linkoping, Nov. 2006. doi:10.1109/NORCHP.2006.329248. © 2006 IEEE. Reprinted with permission.

My Contributions to This Paper

I have written part of the paper and brought forward all the results. This is my first paper on impulse radar solutions exploring CTBV (before we came up with the term). As far as I know this is the first single-chip radar solution reported in standard CMOS technology and is a strong evidence of our new concept. Additional concepts like swept-threshold sampling was introduced and the remarkable performance of the CTBV radar was demonstrated. The constructive use of noise in these systems is also an interesting result.

CMOS Impulse Radar

Håkon A. Hjortland, Dag T. Wisland and Tor Sverre Lande
 Microelectronic Systems, Dept. of Informatics
 University of Oslo, Norway
 Email: haakoh@ifi.uio.no

Claus Limbodal and Kjetil Meisal
 Novelda A.S, Forskningsveien 2A
 0373 Oslo, Norway

Abstract—A novel CMOS Impulse Radar for CMOS implementation is proposed exploring the concept of swept-threshold sampling. Time-domain signal processing with counter-based integration in parallel structures is used. With continuous time, delayline based parallel sampling topologies we achieve a sampling rate in excess of 20GHz. A functional CMOS Impulse radar is implemented in silicon with measured system performance.

I. INTRODUCTION

The idea of impulse radar is nearly a hundred years old. The high frequency pulses or bursts of microwaves are demanding dedicated technology with high frequency devices (SAW filters, bipolars). More recently [1] the impulse radar technique was revived for ground penetrating radar (GPR) for short-range detection of mines buried in the ground. During the nineties McEwan at Lawrence Livermore National Laboratory developed the micropower impulse radar (MIR) for a number of short-range applications preparing ground for simple and power efficient solutions.

The term Ultra Wide Band (UWB) was originally defined as a signal with signal bandwidth $> 25\%$ of the center frequency, but these days UWB is often referred to as the frequency band of [3.1 – 10.6GHz] recently released by FCC. In the following we will use UWB as short for signals exploring the signal band 3.1 – 10.6 GHz as defined by the FCC.

In this paper we are exploring the possibility of short-range radar implementation in standard CMOS technology. Although UWB circuits in non-CMOS technology have been reported [2], to our best knowledge non working single chip solutions in standard CMOS are reported previously. Novel simplified processing techniques sweeping threshold in continuous time in combination with delay lines is taking advantage of technology scaling and is shown to work in silicon.

A. Background

Conventional radar technology is exploring signatures often made up by sinusoidal bursts. The duration of these bursts are limiting the depth resolution of the radar. A shorter signature will increase the resolution. The shortest signature is a single pulse and the shorter pulse length, the better resolution. However, backscattered energy from short pulses are hard to recover and significant signal processing is required.

McEwan [3] proposed the concept of micropower impulse radar (MIR). Randomized pulses are emitted and the backscattered energy sampled at a fixed, but short time interval

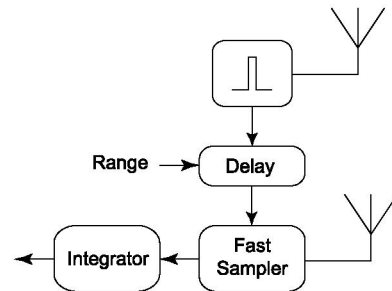


Fig. 1. McEwan micropower impulse radar principle

(strobed) after pulse emission. The integration of backscattered energy is done with an analog integrator and sampling diode(s), preferably fast switching diodes like schottky diodes or step recovery diodes (SRD). The impulse radar may have significant more mileage if available in cheap CMOS technology. However, novel processing techniques and creative use of simple MOS transistors must be utilized to achieve this goal.

As indicated in figure 1, the pulses are emitted repeatedly at randomized intervals and the backscattered energy is sampled after an accurate, fixed delay from pulse emission. A typical pulse duration is $< 1\text{ns}$, maybe close to 100ps, giving wide bandwidth ranging towards 10GHz. The range is set by an accurate delayed strobing after pulse emission. The backscattered signal is weak and often buried in noise. Improved signal-to-noise ratio is achieved by significant integration using fast samplers and analog integrators. An interesting variant of this radar is the motion detecting radar [4] having Doppler-like behavior. Even low frequency movement like heart beat or breathing can be detected as exemplified by Staderini [5] in his demonstrator of a medical radar.

Aiming at a standard CMOS implementation, the high speed signal processing, often analog, is a major challenge. The low power supply voltage virtually prohibiting any kind of high quality analog processing.

The advantage of deep submicron technology is the high speed with gate delay towards 10ps. In this paper we propose novel processing solutions where high quality analog signal processing is simplified to crude thresholding of backscattered energy and continuous time signal processing is explored for signal recovery.

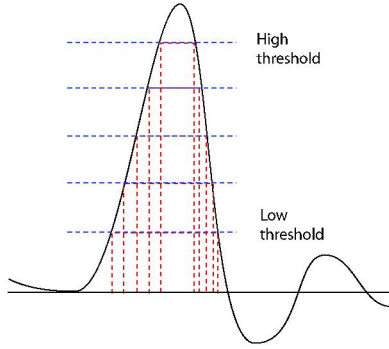


Fig. 2. By sweeping the thresholding level, the incoming signal may be recovered. To some extent the quantized pulse width reflects the strength of the signal.

II. SYSTEM OVERVIEW

The fundamental simplification of using quantized or digital values is well known. In $\Delta - \Sigma$ analog-to-digital converters the analog input signal is converted to a sequence of bits clocked at high frequency known as bitstreams [6]. The pulse density coding (or PWM coding) of bitstreams is accurately representing the continuous analog input value. The coarse thresholding of the continuous input to just '0' and '1' is compensated by increased clock speed. Trading reduced number of quantization levels with higher clock speed is quite appropriate for fine pitch technology and is explaining the gaining popularity of $\Delta - \Sigma$ analog-to-digital converters.

In the CMOS Impulse Radar we are exploring this technique further by processing *two-level quantized (digital) values in continuous time*. We have named this technique "Swept-Threshold sampling" and to our best knowledge sweeping threshold in continuous time is a novel processing technique for impulse radar.

A. Swept-threshold coding

As explained above, the pulses are emitted repeatedly and the backscattered energy accurately strobed. In order to improve signal processing quality we are changing the threshold level "mapping out" the analog input signal as shown in figure 2.

For a given threshold white noise will appear as variations in pulse widths and sometimes even disappearing pulses. By sweeping the threshold we are able to recover quite a lot of the pulse energy and even take advantage of added noise. Pulses below the thresholding level will normally pass undetected but with white noise added, the pulses will sometimes exceed the threshold and the weak signal may be sensed. This behavior is often called *stochastic resonance* [7]. The concept of swept-threshold coding enables quite accurate signal processing even with only two quantization levels.

B. Integration

A crucial function of weak signal reception is significant integration. The purpose of integration is to get rid of added white noise. With repeated pulse emission and thresholding the

white noise will introduce uncertainty in pulse detection. The probability of detecting a pulse, $P('1') = P_{\text{emission}} + P_{\text{noise}}$, will depend on both the signal strength and the noise strength.

In swept threshold coding the noise by itself give a sequence of bits containing an equal number of '1's and '0's (assuming threshold level is in the white noise range), while the emitted imposed signal will contribute with more '1's for positive signals or more '0's for negative signals. The important observation is that the integration can be realized in swept-threshold coding simply by counting '1's and the emitted signal value is the difference from the average value. The integration time is simply the number of pulses emitted and can be controlled. Normally a longer integration time will increase quality, but take longer.

The quality of this style of integration is quite sensitive to the threshold level and behavior is different when the threshold is above the white noise level. With stronger signal levels well above the white noise floor signal detection is simpler. The important sweeping of the threshold level enables iterations to an appropriate threshold level adapting to the received energy of the received pulse.

Analysis of swept-threshold sampling compared to pure analog integration as a function of integration level and noise has been performed (figure 3).

The standard deviation of the white noise term for the analog average improves from σ_N to σ_N/\sqrt{n} by averaging over n samples. With little noise in the input signal, the swept threshold sampling is not doing too well. The surprisingly good performance, however, with increasing noise levels ($\sigma_N = 0.1$ or 1 , 1 being the signal swing), indicates the swept threshold method needs only 10 or even just twice the number of samples of the full analog integrator! These estimates makes the swept-threshold solution quite promising since most real backscattered signal is quite noisy. An interesting observation is that stochastic resonance behavior, only available for large noise components, has the same performance as the swept-threshold sampling.

C. High speed sampling

The simple radar shown in figure 1 is only processing reflecting energy for one strobing time reflecting a fixed distance. Different distances is measured by sweeping the delay and sequentially measure backscattered energy.

With the swept-threshold coding we are able to explore continuous time structures and achieve sampling rates exceeding 20GHz exploring delaylines. The samplers are similar to D-latches with all the data inputs (D) connected to the incoming quantized signal. By clocking (or strobing) the latches in a fast sequence, we are able to accumulate reflected energy at several distances for each emitted pulse.

In figure 4 the principle is shown. We are latching the quantized input pulse with a clock input or strobe pulse delayed by small delays (τ) consisting of two inverters. In the current implementation 64 parallel samplers are used and the inverter delay is approximately 21ps with load measured. In this way we are sampling the incoming pulse with a sampling

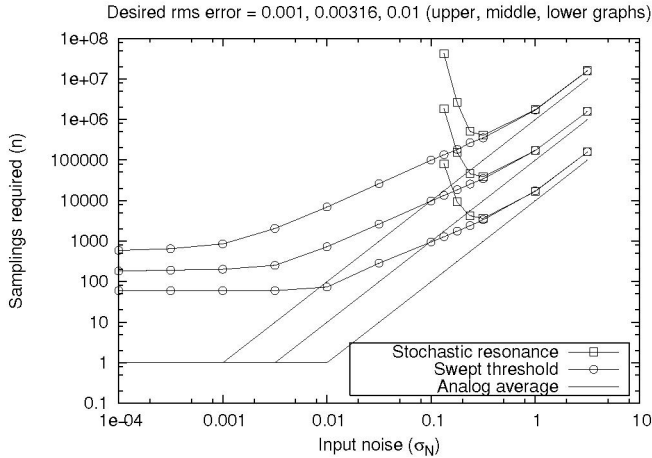


Fig. 3. Comparison of pure analog integration with swept threshold integration indicates a fairly strong integration to perform significantly better. For high noise levels, the swept-threshold technique is doing almost as good as analog integration

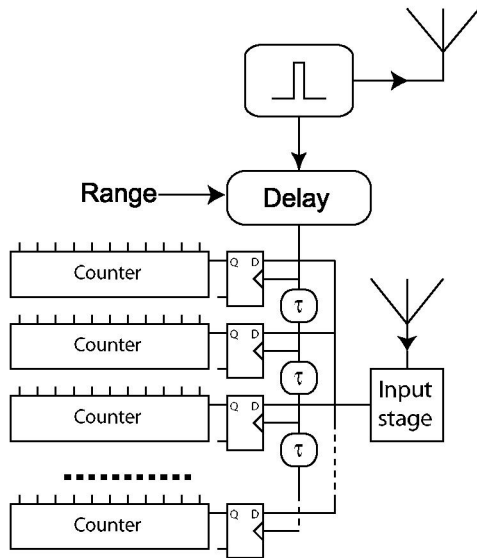


Fig. 4. Delayline sampler and integrators

distance of about 43ps which is equivalent to a sampling rate of $\approx 23GHz$ confirmed by measurements. The depth resolution achieved should be less than 9mm assuming speed of light electromagnetic propagation (in most cases the propagation will be slower, increasing spatial resolution).

With the parallel sampler we are sampling at least 41cm depth. In order to adjust further we have added an initial delay (similar to the McEwan radar) before triggering the sampling sequence. This delay is adjustable, again using delaylines. We have one coarse delayline of $\approx 64 \times 1.4ns$ delay elements and a fine tuning delayline of $64 \times 43ps$ delays. By selecting a coarse and a fine tap using multiplexers we may vary the initial delay from 0 to 90ns with 43ps resolution. With this tuning ability, we are theoretically able to detect targets at a range of up to $> 13m$ (using the speed of light as an estimate

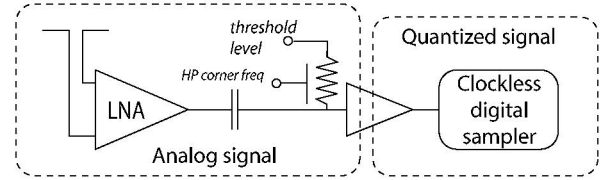


Fig. 5. Input stage of CMOS Impulse Radar consists of an LNA, thresholding element with a buffer amplifier. A simple first order HP filter is included on the input of the thresholding element using a resistive load to the threshold voltage

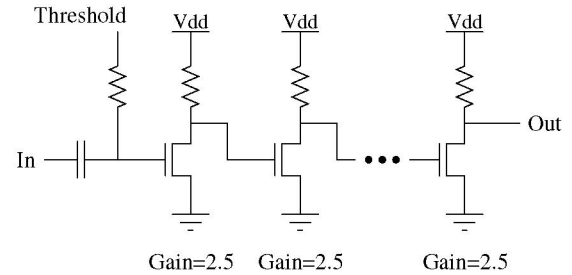


Fig. 6. Common source buffer amplifier consists of eight cascaded common source passive load amplifiers with high-pass tunable filter on the input.

of electromagnetic signal propagation).

As will be shown by measurements all these novel processing techniques are explored to implement a working short range radar in CMOS.

III. IMPLEMENTATION

Based on the principles above we have implemented a prototype of the swept-threshold radar in ST Microelectronics 90nm technology. The process was quite new and the design kit lacked some features for doing high quality layout. It should be noted that this first implementation is merely a conceptual study,

The input stage (figure 5) of the proposed CMOS radar consists of a low-noise amplifier (LNA) designed for a signal bandwidth of 3.1–10.6GHz. Several LNA topologies in CMOS with required bandwidth are published [8] [9]. Typically 10dB gain is achievable with careful design.

The threshold element is AC coupled to the LNA by a blocking capacitor and by driving the input to the buffer amplifier through an equivalent resistance, both thresholding and high pass filtering is achieved. The resistive element is just a weak transconductance amplifier with tunable tail current setting the corner frequency of the HP filter. Some integration due to parasitic capacitance is unavoidable on the input node of the buffer amplifier.

The buffer amplifier consists of eight cascaded common source amplifiers with a total gain of $2.5^8 = 1526$ (figure 6). We are avoiding the high-gain single stage structures like inverters, since they are hard to linearize for the required frequency band. By using several low-gain common-source amplifiers, we are able to maintain an overall linear gain as indicated in figure 7. A total of eight cascaded common-source amplifiers were used.

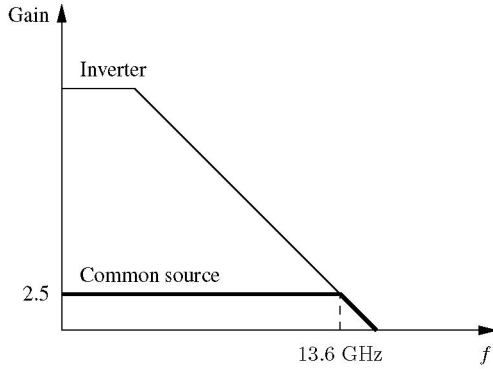


Fig. 7. Buffer amplifiers bode plot indicates the simple idea of controllable flat gain for high frequencies explored in the current prototype.

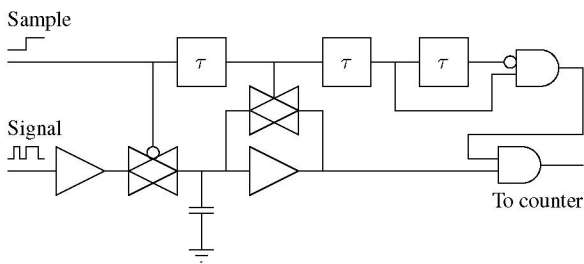


Fig. 8. The sampler consists of several latching elements as the one shown here.

The drawback of this approach is significant power consumption, but with emphasis on novel system aspects, the cascaded solution was chosen.

Leaving the pure analog domain the next stage is sampling of the quantized signal. Since the signal may change rapidly, a specialized latching element has been designed (figure 8). The input signal is driven through an open transmission gate and the input voltage is latched when the transmission gate is strobed. The latched voltage is then shaped by turning on positive feedback with another transmission gate over an amplifier driving the value to '0' or '1'. If the latched value is '1', the pulse duration is shaped to $\approx 1ns$. Suitable timing is achieved inserting inverters. The result of the sampler is a $1ns$ counting pulse for the counter (integrator) when a '1' is detected.

As shown in figure 4 the latching elements are all reading the same input, but the strobing signal is generated by a delayline with two inverter ($\approx 23ps$) between samples.

In order to program the sensing range, a digitally controllable initial delay element is included as explained above

The CMOS impulse radar has been implemented in STMicroelectronics 90nm CMOS technology sharing silicon with other projects (figure 9). Included on the chip was a pulse transmitter designed for UWB signal generation emitting a dual-slope Gaussian pulse of approximately $300mV$ peak-to-peak voltage swing.

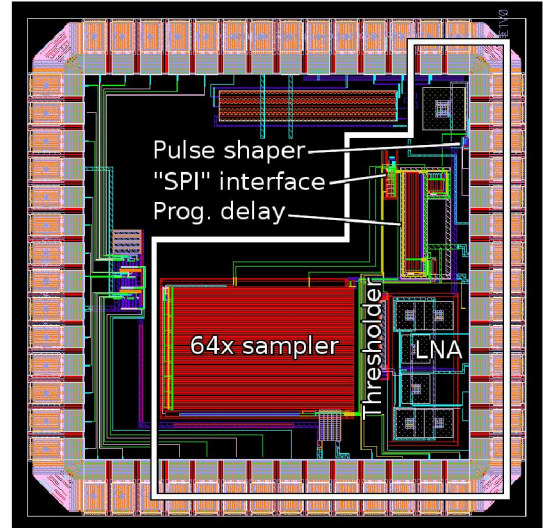


Fig. 9. Layout of the CMOS impulse radar enclosed in the area framed with a white line.

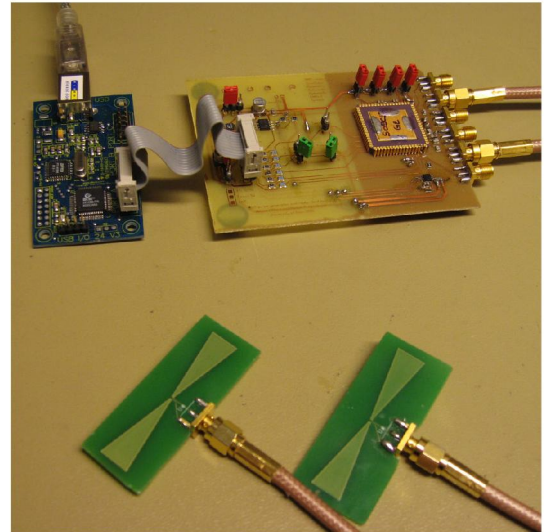


Fig. 10. The measurement setup with the packaged chip mounted on a custom made PCB and interfaced to a PC using a parallel-to-USB interface.

IV. MEASURED RESULTS

The package chip was mounted on a suitable PCB (figure 10) and interfaced through a simple parallel-to-USB card (Elexol USB I/O 24 V3). With suitable software for the connected PC, we could trace movements quite accurately.

In figure 11 the integrator count is indicated by whiter color for high counter values, while black is indicating a low counter values. A hand close to the antennas shows the wave pattern backscattered energy from the hand. When moving the hand away from the radar, the same pattern is shifted rightward as expected. Then again the tracing of moving the hand back again establishing a similar reflection pattern. We are still in an early testing phase, but the first results are already convincing. In our initial experiments we even bypassed the LNA and still get accountable results.

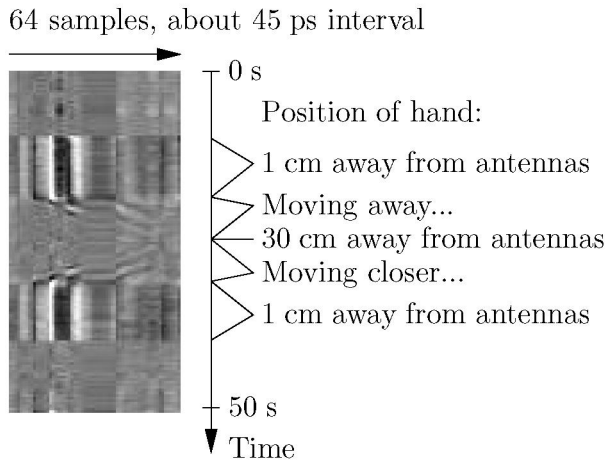


Fig. 11. Measured chip performance moving a hand in front of the radar.

TABLE I
CMOS IMPULSE RADAR SPECIFICATIONS

Measured sampling frequency	23GHz (6.5mm)
Measured initial delay	0–90ns (0–13.5m)
Parallel sampler	64 samples (41cm)
Sampling method	Swept threshold
Signal processing	Quantized continuous time-domain

Through these simple experiments all our assumptions are verified and the implemented CMOS impulse radar is a functional system.

V. CONCLUSION

A working CMOS impulse radar is designed, implemented and tested exploring novel signal processing techniques:

- Binary data sweeping threshold in continuous time (swept-threshold coding).
- Parallel, continuous time high-speed sampler (23GHz measured).

The swept-threshold coding scheme is making implementation in CMOS feasible exploring time-domain processing. Thresholding of input signals simplifies the integration to digital counters. The high speed, parallel sampling structure using delaylines is giving increased depth resolution over a wide programmable range for every emitted pulse. The estimated system characteristics is summarized in table I and the proposed CMOS impulse radar is verified in silicon.

REFERENCES

- [1] D. L. Black, "An overview of impulse radar phenomenon," in *Aerospace and Electronics Conference, 1992. NAECON 1992., Proceedings of the IEEE 1992 National*, vol. 1, Dayton, OH, May 18–22, 1992, pp. 320–326.
- [2] M. Rossberg, J. Sachs, P. Rauschenbach, P. Peyerl, K. Pressel, W. Winkler, and D. Knoll, "11 ghz sige circuits for ultra wideband radar," in *Bipolar/BiCMOS Circuits and Technology Meeting, 2000. Proceedings of the 2000*, Dayton, OH, Sept 24–26 2000, pp. 70–73.
- [3] T. E. McEwan, "Ultra-wideband radar motion sensor," U.S. Patent 5361 070, Nov. 1, 1994.
- [4] —, "Modulated pulse doppler sensor," U.S. Patent 6426716, July 30, 2002.
- [5] E. M. Staderini, "Uwb radars in medicine," *IEEE Aerosp. Electron. Syst. Mag.*, vol. 17, no. 1, pp. 13–18, Jan. 2002.
- [6] S. R. Norsworthy, R. Schreier, and G. C. Temes, *Delta-Sigma Data Converters : Theory, Design, and Simulation*. Wiley-IEEE Press, 1997, no. ISBN: 0780310454.
- [7] D. G. Luchinsky, R. Mannella, P. V. E. McClintock, and N. G. Stocks, "Stochastic resonance in electrical circuits. ii. nonconventional stochastic resonance," *IEEE Trans. Circuits Syst. II [see also Circuits and Systems II: Express Briefs, IEEE Transactions on]*, vol. 46, no. 9, pp. 1215–1224, Sept. 1999.
- [8] A. Bevilacqua and A. M. Niknejad, "An ultra-wideband cmos lna for 3.1 to 10.6 ghz wireless receivers," in *Solid-State Circuits Conference, 2004. Digest of Technical Papers. ISSCC. 2004 IEEE International*, vol. 1, Feb. 15–19, 2004, pp. 382–333.
- [9] K. Meisal, C. Limbodal, T. S. Lande, and D. Wisland, "Cmos impulse radio receiver front-end," in *NORCHIP Conference, 2005. 23rd*, Nov. 21–22, 2005, pp. 133–136.

Paper 2 Thresholded samplers for UWB impulse radar

H. A. Hjortland, D. T. Wisland, T. S. Lande, C. Limbodal, and K. Meisal. “Thresholded samplers for UWB impulse radar”. In: *Circuits and Systems, 2007. ISCAS 2007. IEEE International Symposium on*, pp. 1210–1213, New Orleans, LA, May 2007. doi : 10.1109/ISCAS.2007.378326. © 2007 IEEE. Reprinted with permission.

My Contributions to This Paper

I have written most of this paper and provided all results. Here the CTBV thresholding technique is developed further, exploring some of the qualities of these systems.

Thresholded samplers for UWB impulse radar

Håkon A. Hjortland, Dag T. Wisland and Tor Sverre Lande
Microelectronic Systems, Dept. of Informatics
University of Oslo, Norway
Email: haakoh@ifi.uio.no

Claus Limbodal and Kjetil Meisal
Novelda A.S, Forskningsveien 2A
0373 Oslo, Norway

Abstract— In this paper, we will present two novel methods for sampling the backscatter in an impulse radar system. We have called the two related methods for swept threshold and stochastic resonance sampling. The samplers are simple, mostly digital circuits which are not clocked, but instead utilize continuous-time signal processing. Since fine-pitch CMOS is not very good for analog processing, but instead has very fast digital logic, the samplers are well suited for this technology. An implementation in 90 nm CMOS is described and measurements which confirm a working 23 GHz sampler are shown.

I. INTRODUCTION

The FCC recently released a large unlicensed low power spectrum slot from 3.1 to 10.6 GHz, which opens up for many interesting applications of Ultra-WideBand (UWB) technology. One of these is short range UWB radar, which has a wide but low power spectrum, and thus can be made to fit the FCC emission mask very well. This can easily be accomplished by having the radar transmit very short pulses on the order of a fraction of a nanosecond in length. These short pulses, known as video pulses in radar terminology, will naturally have very wide spectra, and by choosing their exact shape the spectra can also be controlled. Gaussian pulses or their derivatives are good candidates here [1]. The use of low voltage video pulses instead of high voltage carrier-based pulses with coding, can greatly simplify the circuits of a short range UWB radar system compared to traditional radar systems.

There are many possible uses for short range UWB radar systems. At the moment, Ground Penetrating Radars (GPRs) are the most widely used UWB radars [2], but other uses such as heart rate and respiration monitor, through wall sensing and range detection are also very interesting.

A radar works by sending out electromagnetic energy and then observing the backscatter from the targets. In our case, we want to send out a short pulse and then observe the returned signal, which, when there are only a few targets, could be a few discrete, faint echoes of the transmitted pulse. It is possible to do simple detection of for instance the first return pulse to measure distance to the nearest target, but for more advanced purposes, a full readout of the returned signal is necessary, i.e. we need a sampler. Furthermore, since the Signal to Noise Ratio (SNR) drops quickly with distance, we need a system which can integrate several consecutive measurements to create an average, thus increasing the SNR.

Not much work has been done in creating a single chip UWB radar solution, but one chip that has been made is

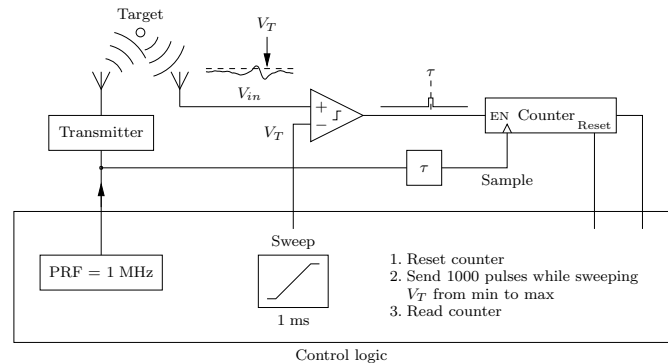


Fig. 1. Swept threshold sampler

described in [3]. This is however done in a SiGe process, whereas the chip in this paper has been made using a standard CMOS process.

II. THRESHOLDED SAMPLERS

The samplers which we will present in this paper are based on known basic principles, but the exact configuration and the use in a UWB radar is probably new.

We will now describe the two samplers, which are very closely related to each other. They can both be realized using the same circuit, so they could perhaps just be called two operating modes of a sampler.

The samplers are mostly based on digital circuitry which is not clocked, but rather uses continuous-time signal processing. As such, they are well suited for fine-pitch CMOS, which is poor at analog signal processing, but in return very fast at digital processing.

The radar system was first presented in [4], and in this paper we describe in more detail how the two samplers work.

A. Swept threshold sampler

The system we have called a swept threshold sampler is shown in figure 1, and details of the sampling process are shown in figure 2. The backscatter, $V_{in}(t)$, which is an ideal $V_{in}(t)$ plus noise, is thresholded with a threshold V_T , and the resulting digital unclocked signal is sampled at time τ after the radar pulse was emitted. To take a single sampling at a given τ is called strobed sampling. By repeating the sampling several times while sweeping the threshold, we can recover the amplitude of the input signal. If, for instance,

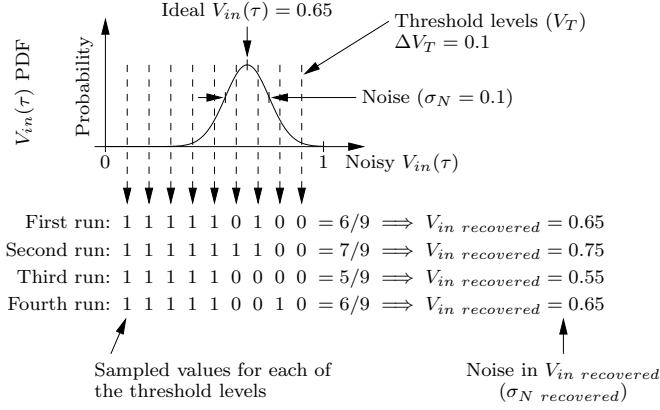


Fig. 2. Swept threshold sampling process

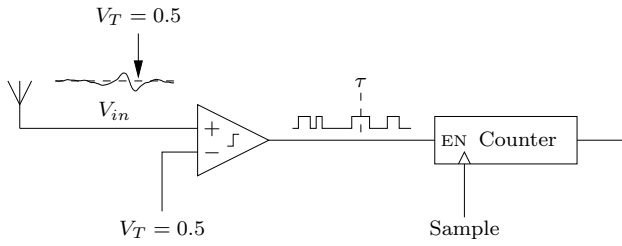


Fig. 3. Stochastic resonance sampler

the ideal $V_{in}(\tau) = 0.65$, and we sweep the threshold from 0.1 to 0.9 V in steps of 0.1 V, we should sample 1's since we are above the threshold, for $V_T = 0.1, \dots, 0.6$, and 0's for $V_T = 0.7, \dots, 0.9$. This means the counter holds a value $c = 6$ after $n = 9$ repeated samplings. These values can then be used to calculate a best guess for the ideal $V_{in}(\tau)$, which we will call $V_{in, recovered}(\tau)$. Since there is noise, σ_N , in the input signal, however, some samples with V_T near 0.65 will be wrong, and we will thus get noise in $V_{in, recovered}(\tau)$, which we will call $\sigma_{N, recovered}$.

The swept threshold sampler thus has many similarities with Analog to Digital Converters (ADCs) such as ramp-compare, flash and delta-sigma, in that many comparisons are made in order to digitize the signal.

B. Stochastic resonance sampler

Stochastic resonance is a phenomenon where signals which have a too low amplitude to pass through a non-linear system, can pass through it anyway if noise is added to the input. The noise will make the weak signal cross some sort of a threshold in the non-linear system from time to time.

The system we have called a stochastic resonance sampler is shown in figure 3, and details of the sampling process are shown in figure 4. The stochastic resonance sampler can be made using the same circuit as the swept threshold sampler, but by just setting the threshold V_T to always be equal to the DC level of the input signal, 0.5 V in this example. If there was no noise, the output of the thresholder would be equal at each

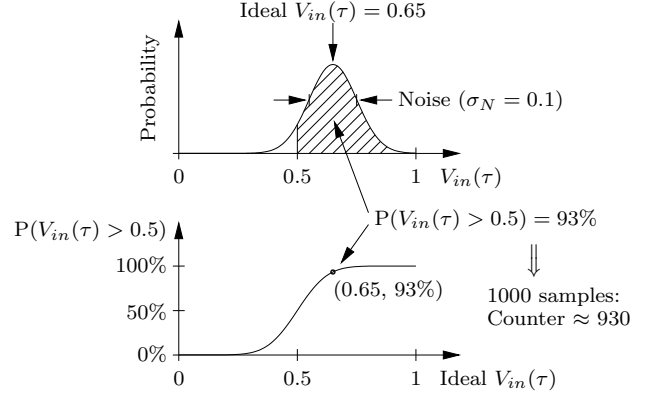


Fig. 4. Stochastic resonance sampling process

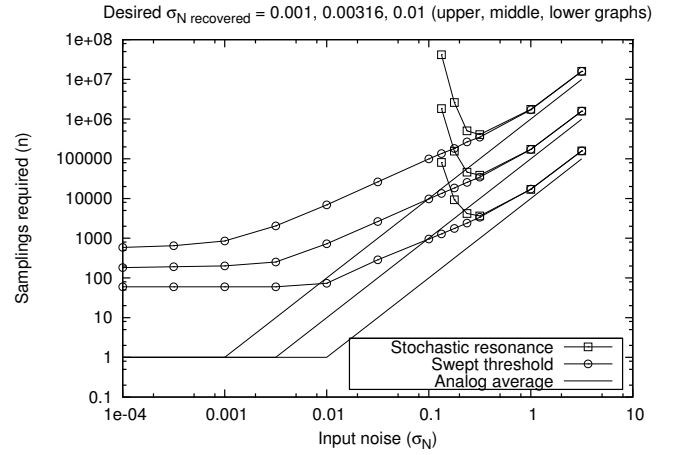


Fig. 5. Simulated performance of samplers

repeated sampling, so this sampler would then only be a simple 1-bit sampler which only detects whether the signal is above or below the DC level. If there is noise, however, $V_{in}(\tau)$ will sometimes be above the threshold and sometimes be below. If for instance the ideal $V_{in}(\tau) = 0.5$, $V_{in}(\tau)$, which is noisy, will be above the threshold 50% of the times. If for instance the ideal $V_{in}(\tau) = 0.65$, the percentage will be higher, and will depend on the amount of noise in the input signal. If the probability of $V_{in}(\tau)$ being over the threshold is for instance 93%, then 1000 samplings will result in a counter value of about 930, but this number is again subject to noise. If we know the Probability Density Function (PDF) of the noise, we can use the counter value to calculate the best guess for the ideal $V_{in}(\tau)$, i.e. $V_{in, recovered}(\tau)$. Again we call the noise in this recovered value for $\sigma_{N, recovered}$.

Note that the stochastic resonance sampler has similarities with the dither method which is sometimes used in traditional ADCs, and also relates very closely to the Suprathreshold Stochastic Resonance (SSR) phenomenon, described in [5].

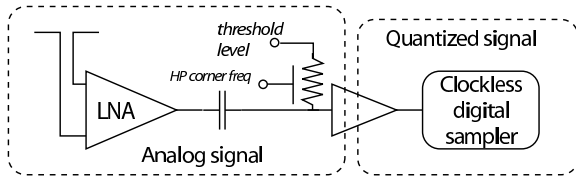


Fig. 6. Input stage

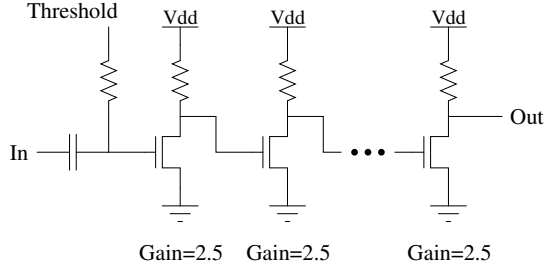


Fig. 7. Common source thresholder

C. Simulation results

A system level simulation of the samplers has been performed, and the result can be seen in figure 5. For comparison purposes we introduce an ideal sampler which we will call the analog average sampler. This is simply a mathematical ideal sampler which calculates the average of all the n samples without thresholding them first. This is probably the best possible sampler or at least very close.

The graph shows how many samples, n , is required in order to achieve a given noise, $\sigma_{N \text{ recovered}}$, in the recovered signal. The big surprise is that the simple thresholded samplers perform very well, i.e. close to the analog average sampler, when there is much noise in the input signal. This means that when sampling noisy signals, these simple samplers might be a good choice.

III. IMPLEMENTATION

Based on the principles above we have implemented a prototype of the swept threshold radar in STMicroelectronics 90 nm technology. The process was quite new and the design kit lacked some features for doing high quality layout.

The input stage (figure 6) of the proposed CMOS radar consists of a low-noise amplifier (LNA) designed for a signal bandwidth of 3.1–10.6 GHz. Several LNA topologies in CMOS with required bandwidth are published [6] [7]. Typically 10 dB gain is achievable with careful design.

The threshold element is AC coupled to the LNA by a blocking capacitor and by driving the input to the thresholder through an equivalent resistance, both thresholding and high pass filtering is achieved. The resistive element is just a weak transconductance amplifier. Some integration due to parasitic capacitance is unavoidable on the input node of the thresholder.

The thresholder consists of eight cascaded common source amplifiers with a total gain of $2.5^8 = 1526$ (figure 7). We are

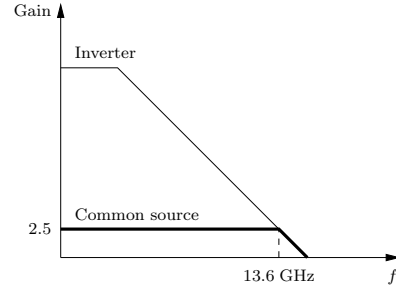


Fig. 8. Common source thresholder vs. inverter frequency characteristic

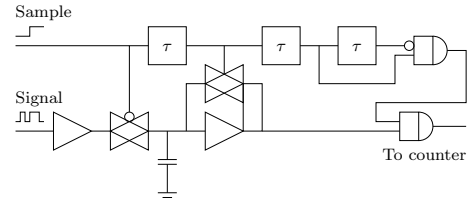


Fig. 9. Digital sampler

avoiding the high-gain single stage structures like inverters, since they do not have a constant gain over the desired frequency band. By using several low-gain common-source amplifiers, we are able to maintain such a constant gain as indicated in figure 8.

The drawback of this approach is significant power consumption, but with emphasis on novel system aspects, the cascaded solution was chosen.

Leaving the pure analog domain the next stage is sampling of the quantized signal. Since the signal may change rapidly, a specialized latching element has been designed (figure 9). The input signal is driven through an open transmission gate and the input voltage is latched when the transmission gate is strobed. The latched voltage is then shaped by turning on positive feedback with another transmission gate over an amplifier driving the value to '0' or '1'. The result of the sampler is a 1 ns counting pulse for the counter (integrator) when a '1' is detected.

Figure 10 shows the $64\times$ sampler implemented on the chip. The latching elements are all reading the same input, but the strobing signal is generated by a delayline with two inverter delays between samples.

In order to program the sensing range, the system also has a digitally controllable initial delay element (τ).

The CMOS impulse radar has been implemented in STMicroelectronics 90 nm CMOS technology sharing silicon with other projects (figure 11). Included on the chip was a pulse transmitter designed for UWB signal generation emitting a first-derivative Gaussian pulse of approximately 300 mV peak-to-peak voltage swing.

A. Measurements

Figure 12 shows a measurement of a moving hand made with the CMOS implementation of the UWB impulse radar,

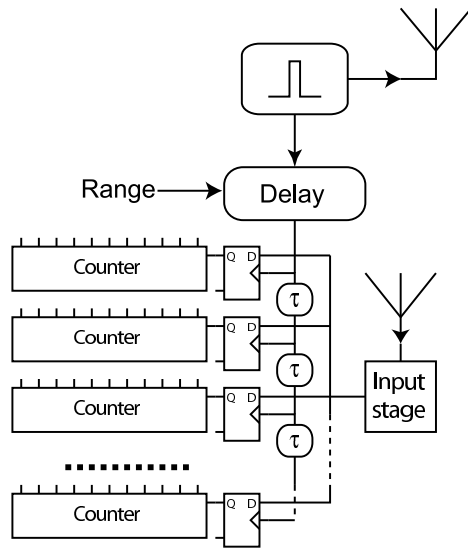


Fig. 10. Delayline sampler and integrators

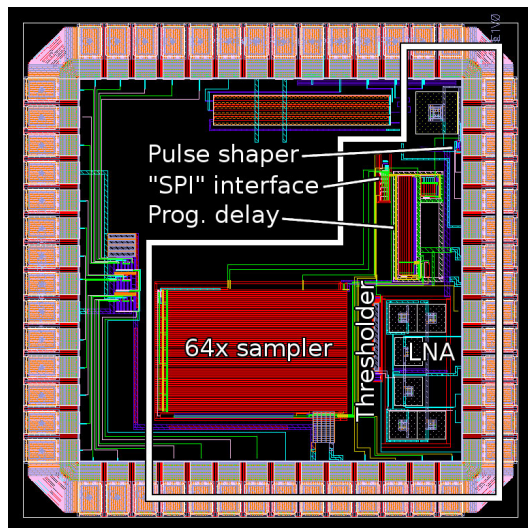
Fig. 11. Layout in 90 nm CMOS. The chip is about 1.2×1.2 mm.

TABLE I

MEASURED CHARACTERISTICS OF CMOS IMPULSE RADAR

PRF	39 kHz (slow meas. equipment)
Transmitted pulse shape	First-derivative Gaussian
Transmitted pulse amplitude	300 mV peak-to-peak
Transmitted pulse center freq.	≈ 2.5 GHz
Measured sampling frequency	23GHz ($\Delta R = 6.5$ mm)
Estimated initial delay range	0–90ns (0–13.5m)
Parallel sampler	64 samples (41cm)
Sampling method	Swept threshold / stochastic resonance
Signal processing	Continuous time quantized amplitude
Power consumption	Between 15 and 60 mA

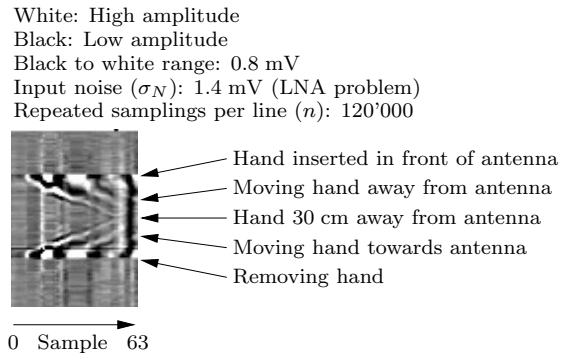


Fig. 12. Measurement of hand movement using the implemented CMOS impulse radar in swept threshold mode

utilizing swept threshold sampling. There were several problems with the implementation, so it does not make much sense with any further, more precise, characterizations of the system. A few more measurements are shown in [8], though. Some of the problems were oscillating LNA, which therefore had to be bypassed, flicker noise in the threshold and interference within the chip.

The sampling rate could be calculated from the figure, but we have done it by using different length loop cables. It was found to be 23 GHz. A summary of the radar characteristics is shown in table I.

IV. CONCLUSION

The swept threshold and stochastic resonance samplers have been presented, and they have been shown through simulation to be able to perform very good for noisy input signals. The samplers are well suited for CMOS implementation, and an implementation in 90 nm CMOS has been described and measurements have been shown, indicating that the samplers also work in actual circuits. The sampling rate has been measured to be 23 GHz. For further details, see [8], which is also the source of most of the figures in this paper.

REFERENCES

- [1] K. Siwiak and D. McKeown, *Ultra-Wideband Radio Technology*, 1st ed. John Wiley & Sons Ltd, 2004.
- [2] A. Yarovoy, "Ultra-wideband systems," in *Microwave Conference, 2003. 33rd European*, 2003.
- [3] M. Rossberg, J. Sachs, P. Rauschenbach, P. Peyerl, K. Pressel, W. Winkler, and D. Knoll, "11 ghz sigc circuits for ultra wideband radar," in *Bipolar/BiCMOS Circuits and Technology Meeting, 2000. Proceedings of the 2000*, 2000.
- [4] H. A. Hjortland, D. T. Wisland, T. S. Lande, C. Limbodal, and K. Meisal, "Cmos impulse radar," in *Proceedings 24th Norchip Conference*, 2006.
- [5] N. G. Stocks, "Suprathreshold stochastic resonance in multilevel threshold systems," *Phys. Rev. Lett.*, vol. 84, no. 11, pp. 2310–2313, Mar 2000.
- [6] A. Bevilacqua and A. M. Niknejad, "An ultra-wideband cmos lna for 3.1 to 10.6 ghz wireless receivers," in *Solid-State Circuits Conference, 2004. Digest of Technical Papers. ISSCC. 2004 IEEE International*, vol. 1, Feb. 15–19, 2004, pp. 382–533.
- [7] K. Meisal, C. Limbodal, T. S. Lande, and D. Wisland, "Cmos impulse radio receiver front-end," in *NORCHIP Conference, 2005. 23rd*, Nov. 21–22, 2005, pp. 133–136.
- [8] H. A. Hjortland, "Uwb impulse radar in 90 nm cmos," Master's thesis, Univ. of Oslo, 2006. [Online]. Available: <http://urn.nb.no/URN:NBN:no-12777>

Paper 3 Impulse Radio technology for Biomedical applications

T. S. Lande and H. A. Hjortland. “Impulse Radio technology for Biomedical applications”. In: *Biomedical Circuits and Systems Conference, 2007. BIOCAS 2007. IEEE*, pp. 67–70, Montreal, Que., Nov. 2007. doi : 10.1109/BIOCAS.2007.4463310. © 2007 IEEE. Reprinted with permission.

My Contributions to This Paper

This paper was written jointly with my supervisor and my contributions are the radar part.

Impulse Radio technology for Biomedical applications

Tor Sverre Lande and Håkon A. Hjortland
 Nanoelectronics, Dept. of Informatics
 University of Oslo, Norway
 Email: bassen@ifi.uio.no

Abstract—Improving quality of service in wireless communication links is of vital importance in biomedical applications. Current wireless technology is far from satisfactory conveying vital signs in personalized healthcare. Nevertheless current technology like Bluetooth and ZigBee is explored even for critical monitoring of patients, both in hospital environments and homecare. In this paper we are exploring impulse radio as a feasible technology for health monitoring. By exploring advanced technology and novel architectures, improved quality of service may be granted. Additional interesting biomedical functionality of impulse radio is detached body sensors (short-range medical radar).

I. INTRODUCTION

As wireless functionality is applied in an increased number of daily life functions, most of us are accepting the fact that connectivity is not always present. We just move around a little or try again later. We are even relying on wireless connectivity of cellular phones in critical situations. Although many wireless services have extensive coverage, the quality of service (QoS) is currently not at the level of guaranteed connectivity. In addition to coverage, traffic load due to limited number of channels and noise interference is also affecting service quality. Apparently the benefit of getting wireless is greater than the risk of losing information.

In a professional health environment like a hospital, are we ready to accept the quality of service provided of available wireless technology? Can life support equipment be wireless? In a fairly static situation (patient in bed) quite good connectivity may be offered by current technology like Bluetooth or ZigBee, but the benefits of wireless is mostly appreciated in a dynamic environment where patients are moving around. Even with extensive in-house coverage, can connectivity be guaranteed? Would current technology provide connectivity through the body of a fainted patient? These questions are getting even more pressing when hospital services are provided remotely at home. As the number of elderly people is growing, wireless personalized healthcare in home environment is beneficial both for the patient and for the healthcare provider with reduced costs. However, current wireless technology is not by far good enough for critical monitoring of vital signs in a dynamic situation.

II. TECHNOLOGY

Are there any technological solutions providing the robustness of operation required for health monitoring? In the

following we will try to identify bottlenecks and possible solutions.

- 1) **Sensitivity.** An obvious improvement would be increased sensitivity of wireless technology. Carrier based, narrow-band receiver solutions available already have impressive performance with sensitivity approaching -100 dBm. Although improvements may appear, the increased sensitivity will only be marginal.
- 2) **Capacity.** Capacity in wireless networking is often measured in available transmission bandwidth. Contrary to computer networking, health monitoring does not require large transmission bandwidth. Most health readings are low bandwidth. Typically glucose measurements only require a handful of readings each hour! One of the most demanding monitoring is ECG. Typically 12 bit resolution at 250 Hz sample rate is used giving a data rate of only 3 kbit/s.
- 3) **Penetration.** A major issue in health networks is signal penetration. RF signals are heavily attenuated by the body (mostly fluids) and penetration is inverse proportional to the carrier frequency. Studies [1] [2] indicate very little energy is penetrating body tissue at microwave narrowband frequencies.
- 4) **Channels.** In health networks used in hospitals crowded with patients, the number of available channels and channel interference is a limiting factor. Although time-multiplexing using TDMA/CDMA solutions in combination with cellular structure enable larger number of simultaneous transmitters, quite complex infrastructure and overhead is necessary to maintain good response.
- 5) **Transmission power.** In order to maintain connectivity carrier based radio transmission is burning quite some power during data transmission. In combination with complex networking protocols, battery life of portable devices is limited.

There are no final and absolute solutions to the limitations of wireless technology, but we will explore impulse radio technology to improve on some of these issues.

III. IMPULSE RADIO

The concept of impulse radio is the oldest communication technology explored by Marconi in 1901 using wide band sparks for wireless communication. The wideband “sparky” transmission was soon replaced with narrowband, carrier based

solutions still dominating wireless technology. However, US regulating authority (FCC) recently released the frequencies 3.1–10.6 GHz for unlicensed use provided emission levels are kept low (< -41.3 dBm/MHz). This new unlicensed band, often called Ultra-WideBand (UWB), is the largest unlicensed frequency band ever released. Although the prime motivation for releasing this frequency band was high speed wireless communication, the UWB frequencies may be explored in different ways gaining other advantages in health applications. The released frequency band is wide enough for proper impulse radio applications. In the following we will use the term Ultra-WideBand Impulse Radio (UWB-IR) to denote impulse radio transmission for the 3.1–10.6 GHz band.

A noticeable exploration of impulse radio technology is the work by Tom McEwan on Micropower Impulse Radar (MIR) [3] [4]. Small energy bursts are emitted as impulses (noise) and the very weak reflected energy is recovered using range gating in combination with significant integration (analog). Staderini [5] was able to recover heart beats remotely exploring MIR architecture. Two important observations of these works are 1) the penetration of heavy materials with pulses and 2) the ability to recover very small amounts of reflected energies using integration. Both these features are desirable also from a communication perspective.

A major issue of the MIR technique is the fundamental analog nature in combination with the high speed sampling required. Implementation of MIR in low voltage CMOS technology is simply not possible and to our best knowledge we have the only CMOS implementation of impulse radar [6].

Fundamental new signal processing techniques is required to facilitate the wideband processing and high speed sampling required for UWB-IR. In our work “Thresholded samplers for UWB impulse radar” [6] we are presenting different coding strategies and technology friendly signal processing techniques suitable for energy efficient impulse radio implementations in CMOS. The main features are:

- **Thresholding** of incoming signal giving binary, single bit signal. The threshold level may be varied through repetition resembling different quantization levels. In [6] the concept of “swept threshold sampling” is elaborated.
- **Continuous time** coding of signals. No clocks are used enabling a tenfold increase in sampling frequency.

Modern technology is designed for digital signal encoding and by avoiding clocks, speed is improved and power efficiency is increased significantly.

In order to show the benefit of this new approach, our impulse radar architecture is shown.

IV. SWEEPED THRESHOLD RADAR

The swept threshold radar [6] works by transmitting UWB impulses, then thresholding and sampling the received echo signal. The digital but unclocked output from the thresholder is sampled in sequence by a number of parallel D-flip-flops, all having the thresholder output as their inputs. The D-flip-flops are set up to trigger one after another in rapid succession, with two inverters used as a delay element between them. When the

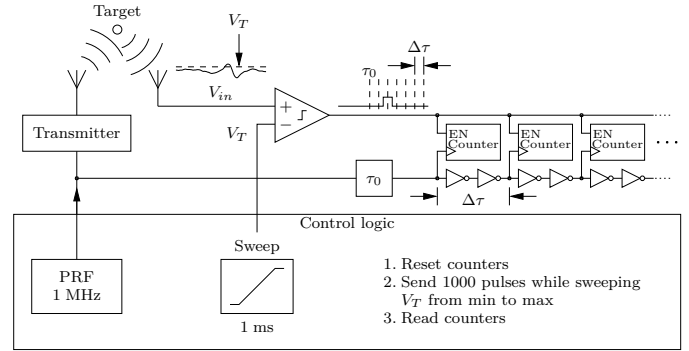


Fig. 1. Swept threshold sampler for radar

signal has been sampled, a counter connected to each of the D-flip-flops is incremented if a “1” was sampled.

Another, more high-level, way of looking at this architecture, is to note the integration action of the latch-and-count structure. As shown in figure 1 these integrators are clocked at high speed ensuring simultaneous sampling at different depths.

By repeating this procedure while sweeping the threshold level from below to above the range of the input signal, a full readout of the received signal is done. The integrating action is achieved through repetition since the sampling is precisely timed. A counter which is clocked for instance at a high peak of the input signal, will sample many “1”s since the signal is above the threshold most of the time. A counter clocked at a low peak of the input signal, will sample few “1”s since the threshold quickly crosses the level of the low peak and the threshold output thus will be “0” for that low peak for most of the sweep. All in all, this architecture samples the received signal completely, which is assumed not to change too much during the sampling, at a high sampling rate given by the delay between the clocking of two adjacent counters. By removing the clutter of static objects, for instance by high-pass filtering the entire readout-curve in post-processing, we get a radar readout showing movements in front of the antennas.

An important detail is the integration of the white noise which always gets added to any received RF signal. This does mean that the counter values will be a little off from their expected values, and when repeating the sampling, the counter values will jump slightly up and down. By increasing the number of pulses sent during the sweep, the noise can be averaged away, or at least reduced to an acceptable level. An ideal analog averaging sampler would have the noise reduced according to

$$\sigma_{N \text{ recovered}} = \frac{\sigma_N}{\sqrt{n}}, \quad (1)$$

where σ_N is the standard deviation of the noise in the input signal, n is the number of repeated samplings and $\sigma_{N \text{ recovered}}$ is the standard deviation of the noise in the recovered signal, i.e. after averaging. We can see that by for instance increasing the number of repeated samplings by a factor of 100, the amplitude of the noise in the readout will be reduced by a factor of 10. Although this is the behavior of an ideal analog

average sampler, the swept threshold sampler actually behaves quite similar, and for noisy signals its performance is even quite close to the performance of the ideal analog average sampler.

This way of sampling signals, by early thresholding and subsequent signal processing on digital, 1-bit, unclocked signals, is particularly well suited for standard fine-pitch CMOS, which is very fast and also relatively cheap. The fact that fine-pitch CMOS with its low supply voltage is not very well suited for analog signal processing is not a problem because we do most of the signal processing in the digital domain. The “digital” integrators (counters) are area efficient. Simultaneous sampling at a large number of different depths is easy to achieve by simply duplicating the integrator and sample with delayed clock. The McEwan MIR only samples a single depth, which of course can be shifted back and forth to get a full readout, but this procedure is time consuming. Since the sampling rate is only determined by the delay between the triggering of the different counters, it can be very high, resulting in improved depth resolution. By using two inverters as a delay element in 90 nm CMOS, we have achieved approximately 30 GHz sampling rate, equivalent to a depth resolution of approximately 5 mm. Another important property of the digital integrator is the integration time. Unlike a lossy analog integrator, the “digital” integrator is lossless and when appropriate, the integration time may be arbitrarily long.

In summary, the swept threshold radar is well suited for fine-pitch CMOS and allows high-speed, multi-point sampling of faint radar echo signals, with noise amplitude proportional to $1/\sqrt{n}$, allowing for arbitrarily high SNR by choosing sufficiently long integration time.

V. IMPULSE RADIO COMMUNICATION

How can energy efficient impulse radio communication solutions be made in CMOS exploring the swept threshold radar technique and what are the benefits? Several companies like Time Domain Solutions have been using impulse radio transmission for communication. Another interesting effort is the low-power definition of the ZigBee physical layer known as IEEE 802.15.4a. One of the proposed radio solutions is impulse based.

An impulse based radio channel is made by transmitting symbols coded as Pseudo Noise (PN) sequences of pulses. The noise-like property is important for smoothing out the spectral content. In general it is hard to avoid spectral peaks and several coding schemes are feasible including biphasic coding. From an implementation perspective transmission of symbols as PN sequences is simple requiring no oscillators for carrier generation. These carrier free transmitters can also be very power efficient. The load is shifted to the receiver demanding running cross correlation looking for expected PN sequences. These receivers are often called correlating receivers and are known to be power hungry and difficult to make.

One major issue is synchronization. Incorporating the clock in the PN sequence will most certainly give undesirable spectral peaks. Aiming for UWB-IR solutions, sampling rates

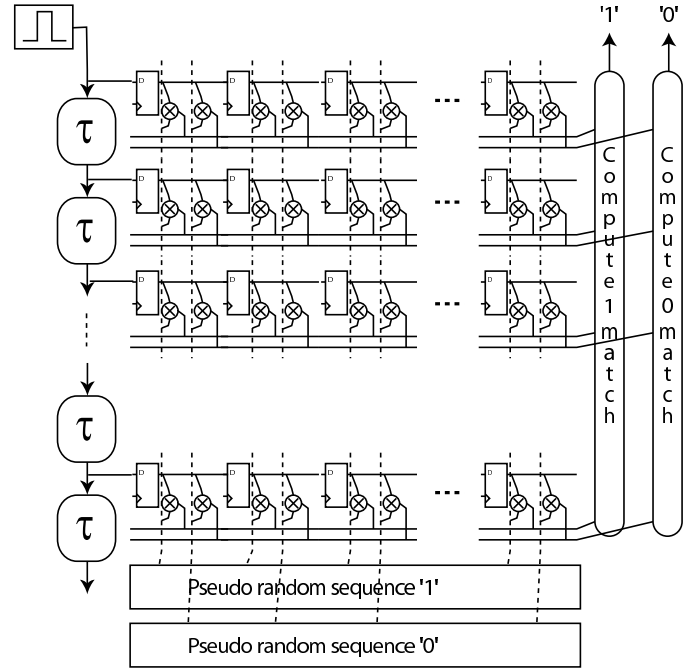


Fig. 2. The thresholding, continuous time full RAKE receiver

at several GHz is hard to make and for sure power consuming. Again, by exploring the concepts of thresholding and continuous time signals we are able to find a power efficient architecture avoiding precise clock synchronization.

The concept of RAKE architecture was proposed a long time ago [7] and is sifting through the incoming signal for PN codes looking for both Line-Of-Sight (LOS) and Non-Line-Of-Sight (NLOS) PN patterns. This pure time-domain processing structure is equivalent to running cross-correlation with some expected template. With short pulses, patterns due to multi-path transmission may also be recovered and used constructively for symbol detection. The LOS PN and the NLOS PNs are usually detected in RAKE fingers. The RAKE function is often implemented as software on DSPs with typically 2–3 RAKE fingers known as partial RAKE receivers.

The continuous time, thresholding RAKE receiver architecture [8] is shown in figure 2. The incoming RF signal is immediately thresholded using a comparator. By feeding the value after thresholding (digital bit) through a delayline using cascaded digital inverters, the pattern of LOS components and NLOS components spanning one chip of the PN sequence may be stored with suitable resolution. Sampling the delayline with the chip clock (typically in the MHz range) we are able to latch the signal value as a bit in a clocked shift register. Since we do not require any clock synchronization, the LOS PN pattern may occur in any of the RAKE fingers constituted by the shift registers. In the end we really do not care which one, since we are doing them all. Correlating bit-patterns is also simple. The multiplications of cross-correlation are reduced to simple AND gates. The clock accuracy is only required for the duration of one PN coded symbol.

So by using thresholding and continuous time signals, we are able to design a power efficient impulse radio, correlating receiver exploring a full RAKE architecture.

The impulse radio transmission scheme has a number of interesting features.

A. Penetration

Several studies [9] indicate the improved penetration of impulses in heterogeneous matter. This is demonstrated in military applications for Ground Penetrating Radar (GPR). Studies using impulse radio on breast cancer detection are also promising [10]. The Staderini medical radar experiments confirm penetration through measurement of heart beat using impulses.

Impulse radio transmission is obeying the same physical laws as standard radio. The higher penetration is due to the wide spectral content of impulses in combination with heterogeneous, often layered, structures found in nature. For a given dimension (thickness of a layer) it is possible to find some frequency with higher penetration due to resonance (matching wavelength). Impulses are characterized by a wide frequency content possibly spanning the available 3.1–10.6 GHz frequency band. With UWB-IR it is quite easy to hit one or more resonating frequencies penetrating quite heavy material. With narrow band, carrier-based transmissions one primary carrier frequency is transmitted with little probability of hitting resonance frequencies.

B. Constructive Interference

As mentioned above unlike fading in carrier based systems multi-path interference may be constructive in UWB-IR systems. If proper orthogonal sequences are chosen, very little interference is occurring, even with a large number of channels.

C. Improved sensitivity

One of the major improvements for robust communication is increased sensitivity. The “integration” action of repeated pulse transmission of the swept threshold radar is equivalent to increasing number of pulses of the PN codes. By using longer PN sequences we may relax the detection level of the RAKE receiver effectively accepting more pulse loss. Evidently longer PN sequences will reduce the transmission

bandwidth, but we are gaining sensitivity. To what extent this technique will outperform the sensitivity of carrier based receivers remains to be seen, but the potential is certainly there.

D. Channelization

It is possible to tune an UWB-IR transmission for a large number of channels (simultaneous transmissions). If sparsely populated PN sequences are used, a large number of orthogonal PN sequences are available enabling a large number of channels (only two symbols are required for each channel). In excess of 100 channels are easy to get with minor interference. Compared to typically 16 channels in narrowband systems, UWB-IR is adapting to densely populated environments (like hospitals).

VI. CONCLUSION

The potential of UWB-IR in biomedical communication is certainly promising. In this paper we have shown the fundamental opportunities of UWB-IR and how technology friendly solutions exist. We are convinced that future implementations will meet the demands of biomedical applications.

REFERENCES

- [1] C. Gabriel and S. Gabriel, “Compilation of the dielectric properties of body tissues at rf and microwave frequencies,” Physics Department, King’s College London, 1996.
- [2] T. Zasowski, G. Meyer, F. Althaus, and A. Wittneben, “Propagation effects in uwb body area networks,” in *Ultra-Wideband, 2005. ICU 2005. 2005 IEEE International Conference on*, 2005, pp. 16–21.
- [3] T. E. McEwan, “Ultra-wideband radar motion sensor,” U.S. Patent 5 361 070, Nov. 1, 1994.
- [4] S. Azevedo and T. E. McEwan, “Micropower impulse radar,” *Science & Technology Review*, January/February 1996.
- [5] E. M. Staderini, “Uwb radars in medicine,” *IEEE Aerosp. Electron. Syst. Mag.*, vol. 17, no. 1, pp. 13–18, Jan. 2002.
- [6] H. A. Hjortland, D. T. Wisland, T. S. Lande, C. Limbodal, and K. Meisal, “Thresholded samplers for uwb impulse radar,” in *Circuits and Systems, 2007. ISCAS 2007. IEEE International Symposium on*, 2007.
- [7] R. Price and P. Green, “A communication technique for multipath channels,” in *Proceedings of the IRE*, 1958, 1958.
- [8] C. Limbodal, K. Meisal, T. S. Lande, and D. Wisland, “A spatial rake-receiver for real-time uwb-ir applications,” in *2005 IEEE International Conference on Ultra Wideband IEEE Cat. No.05EX1170C*, 2005.
- [9] R. M. Narayanan, X. Xu, and J. A. Henning, “Radar penetration imaging using ultra-wideband (uwb) random noise waveforms,” in *Radar, Sonar and Navigation, IEE Proceedings -*, 2004.
- [10] B. Zhou, W. Shao, and G. Wang, “On the resolution of uwb microwave imaging of tumors in random breast tissue,” in *Antennas and Propagation Society International Symposium, 2005 IEEE*, 2005.

Paper 4 CTBV Integrated Impulse Radio Design for Biomedical Applications

H. A. Hjortland and T. S. Lande. “CTBV Integrated Impulse Radio Design for Biomedical Applications”. *Biomedical Circuits and Systems, IEEE Transactions on*, Vol. 3, No. 2, pp. 79–88, Apr. 2009. doi:10.1109/TBCAS.2009.2014962. © 2009 IEEE. Reprinted with permission.

My Contributions to This Paper

This is the main paper about CTBV signal processing and is my major scientific contribution. I have written most of the publication and provided all results while the fundamental ideas were worked out together with my supervisor. The CTBV concept was considered to be new design paradigm when writing the paper and has given remarkable results regarding power-efficient, high-speed solutions. Afterwards we have discovered Y. Tsividis has explored a similar concept (CTDA) in signal processing.

CTBV Integrated Impulse Radio Design for Biomedical Applications

Håkon A. Hjortland and Tor Sverre “Bassen” Lande, *Senior Member, IEEE*

Abstract—Improving quality of service in wireless communication links is of vital importance in biomedical applications. Limitations of current technology are evident with a limited number of channels and prone to fading. In this paper, we are exploring impulse radio as a feasible technology for health monitoring and even as novel detached sensors. By exploring advanced deep submicron technology and novel architectures, improved quality of service may be granted. Additional interesting biomedical functionality of the impulse radio is detached body sensors (short-range medical radar).

Index Terms—Biomedical, communication, complementary metal–oxide semiconductor (CMOS), radar, RAKE, sensor, ultra-wideband–impulse radio (UWB–IR), wireless.

I. INTRODUCTION

THE rapid assimilation of mobile communication is striking and full coverage is assumed. Hiking in the wilderness with the mobile phone as the only security device is quite common. Both limited coverage and battery-life impose strong limitations on reliability. Although many wireless services have extensive coverage, the quality of service (QoS) is currently not at the level of guaranteed connectivity. In addition to coverage, traffic load due to a limited number of channels and noise interference are also affecting service quality. Apparently, the benefit of getting wireless is greater than the risk of losing information.

As the number of elderly people grows, wireless personalized health care in the home environment is beneficial for the patient and for the healthcare provider with reduced costs. However, current wireless technology is not, by far, good enough for critical monitoring of vital signs in a dynamic situation. In the following sections, we will explore wideband impulse radio using the recently released 7.5-GHz wide frequency band from 3.1 GHz to 10.6 GHz. Although strongly limited in power, we believe several interesting features may arise, provided implementations are available in standard complementary metal–oxide semiconductor (CMOS) technology.

II. FEASIBLE TECHNOLOGY

Several narrowband wireless solutions exist. We will nail down some of the important issues and point at some feasible solutions that are adapted for nanometer technology.

Manuscript received April 02, 2008; revised June 27, 2008, October 15, 2008, and January 16, 2009. Current version published March 25, 2009. This paper was recommended by Associate Editor Ralph Etienne-Cummings.

The authors are with the Department of Informatics, University of Oslo, Oslo N-0316, Norway (e-mail: haakoh@ifi.uio.no; bassen@ifi.uio.no).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TBCAS.2009.2014962

- 1) Sensitivity. An obvious improvement would be increased sensitivity of wireless technology. Carrier based, narrow-band receiver solutions available already have impressive performance with sensitivity approaching -100 dBm. Although improvements may appear, the increased sensitivity will only be marginal.
- 2) Capacity. Unlike consumer electronics demands, health applications do not require a very high data rate. For example, monitoring of glucose for diabetes treatment requires less than ten readings per hour.
- 3) Channels. In health networks used in hospitals crowded with patients, the number of available channels and channel interference is a limiting factor. Although time-multiplexing using time-division multiple access (TDMA)/code-division multiple access (CDMA) solutions in combination with cellular structures enables a larger number of simultaneous transmitters, a rather complex infrastructure and overhead are necessary to maintain good response.
- 4) Transmission power. In order to maintain connectivity, carrier-based radio transmission is burning some power during data transmission. In combination with complex networking protocols, battery life of portable devices is limited. This issue is of very high importance for body implants where very little heating is tolerated.
- 5) Interference and multipath fading. Destructive interference or fading is a major issue with carrier-based wireless systems. The center frequency of each channel must be carefully selected and channel separation must be significant to avoid channel interference. In highly reflective environments (indoor), multipath reflection may mix with the direct path signal and attenuate the transmitted signal significantly (fading).
- 6) Localization. As inroom or inhouse wireless technology is adopted for professional and home use, additional functionality is required. The concept of wireless sensor networks (WSN) explores a large number of small wireless nodes for environmental sensing and control [1]. When network nodes are moving around, accurate localization is often of vital importance. Current wireless technology provides low-precision ranging measurements using received signal strength indication (RSSI) with a couple of meters precision.

In the following text, we will explore impulse radio solutions and indicate how these limitations may be circumvented. We are committed to solutions that are implementable in nanometer CMOS technology.

III. IMPULSE RADIO

The concept of impulse radio is the oldest communication technology explored by Marconi in 1901 using wideband sparks for wireless communication. The wideband “sparky” transmission was soon replaced with narrowband, carrier-based solutions still dominating wireless technology. However, the U.S. regulating authority (FCC) recently released the frequencies 3.1–10.6 GHz for unlicensed use provided that emission levels are kept low (< -41.3 dBm/MHz). This new unlicensed band, often called ultrawideband (UWB), is the largest unlicensed frequency band ever released. The released frequency band is wide enough for proper impulse radio applications. In the following text, we will use the term UWB–impulse radio (UWB–IR) to denote impulse radio transmission for the 3.1–10.6 GHz band. Some reported results of using the impulse-based systems may be found in [2]–[4].

A noticeable exploration of impulse radio technology is the work by Tom McEwan on micropower impulse radar (MIR) [5], [6]. Small energy bursts are emitted as impulses (noise) and the very weak reflected energy is recovered using range gating in combination with significant integration (analog). Staderini [4], [7] was able to recover heart beats remotely by exploring the MIR architecture. Two important observations of these works are: 1) the penetration of heavy materials with pulses and 2) the ability to recover very small amounts of reflected energies using integration. Both of these features are desirable also from a communication perspective.

A major issue of the MIR technique is the fundamental analog nature in combination with the high-speed sampling required. The implementation of MIR in low-voltage CMOS technology is simply not possible and to our best knowledge, we have the only CMOS implementation of impulse radar [8].

IV. CTBV SIGNAL PROCESSING

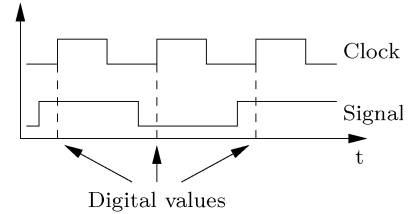
To implement systems exploiting the beneficial properties of UWB–IR, we need circuits which can provide the multi-gigahertz signal processing capabilities required to do so. Not only do we need to perform analog signal processing at these high frequencies, but also nonlinear operations, such as pattern recognition to detect pulses and high-speed sampling to make radar readouts [9]. In this paper, we will use the term continuous-time binary value (CTBV) signal processing, which explores clockless, continuous time use of binary logic in standard CMOS.

Signal-processing domains may be classified into four categories according to whether time and value are quantized or not (Table I). The upper right one, where the value is binary but time is kept continuous, is what we have called CTBV and which we will now explore. We will show that this is a signal-processing domain of its own with a unique set of properties which combines the speed of analog signal processing and some of the advanced processing capabilities available with digital signal processing (DSP). It may not be a replacement for either of the other signal-processing domains, but it will be able to perform some tasks at high speeds which the other domains are not capable of. Another important property is the avoidance of high-speed, power-consuming clocks.

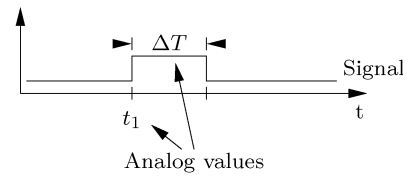
TABLE I
SIGNAL-PROCESSING DOMAINS

Value	Time	
	Discrete	Continuous
Binary	Digital	CTBV
Continuous	Analog sampled-data (switch-cap...)	Analog

Clocked digital signal:



CTBV signal:



Much more information content

Fig. 1. Clocked digital versus CTBV signal representation.

A. Signal Representation

The signals in the CTBV domain are represented somewhat different than in any of the other signal-processing domains. Each signal is a single-bit binary value which is unclocked (i.e., a signal which can be either “0” or “1”), with transitions at any point in time. The data in the signal are thus represented primarily by the transitions rather than the actual binary value (Fig. 1). Such signal representation is already widely in use in coding strategies, such as pulsewidth modulation (PWM) and pulse-position modulation (PPM). In contrast to regular clocked digital logic, we see that the CTBV signal representation holds data of a somewhat analog quality since the transitions do not have any discrete steps. This means that the data conveyed by CTBV signals are analog values at discrete points in time.

As a simple example of CTBV signal representation, consider a pulse with a width of 1 to 2 ns. It could hold an analog value represented by the width of the pulse. Without phase noise in the system, this analog value would have infinite precision. When sending such a signal through an inverter, the inverter will only add a constant, tiny delay to the time of the transition events and, thus, the information represented by the timing of these events is preserved, only a bit skewed in time. Some noise will be added by passing the signal through the inverter and in this signal coding, our source of noise is therefore the phase noise of the logic gates.

B. Processing Operations

CTBV signals can be processed with some specialized operations by using simple circuits made of mostly digital-logic gates. We will now present a few of the basic signal-processing operations which we are using for CTBV-coded signals.

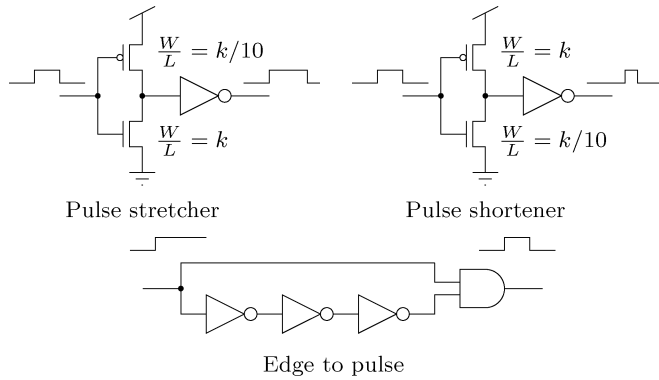


Fig. 2. Examples of CTBV pulse shapers.

1) *Delay*: Two cascaded inverters provide a simple circuit for delaying a CTBV signal. The width of pulses will be preserved through this circuit, so that the signal will not be distorted, only delayed by a given amount.

In 90-nm CMOS technology, two inverters will provide about 30–40-ps delay, and if even shorter delays are desired, it is possible to delay signal “A” with, for example, 35 ps and signal “B” with 40 ps, which gives a relative delay of only 5 ps. In such a way, fairly high relative temporal resolution is feasible.

By cascading several delay elements, we obtain a “delay line,” which provides a memory or, perhaps more precisely, a history of the input signal. For signal processing, this kind of signal history is often very useful. By sampling intermediate states of the delay line, or by using the delay line to trig a sequence of samplers, we may design a system which performs downconversion of signals from a tens of gigahertz sampling rate to rates which are much more manageable by conventional digital logic. This simple, yet very useful architecture has been exploited in the radar and the RAKE receiver described in this paper.

2) *Logic*: Simple combinational logic operations can be applied to CTBV signals. This will detect coincidental states of two or more input signals, for example, when both “A” and “B” are high, whose meaning depends on the exact coding of the input signals. A more interesting example is to combine logic with delay elements, which enables detection of patterns in time, for example, to detect the occurrence of events, such as “A” going high 100 ps before “B.”

3) *Pulse Shaping*: By using asymmetric sizing of the NMOS/PMOS devices in inverters (Fig. 2), it is possible to create circuits which delay either just the rising or the falling edges of signals going through them. In this way, we can easily increase or decrease the width of pulses by a predetermined amount. It is also very simple to create edge detectors which detect, for instance, rising edges and create a fixed-width pulse for each rising edge it encounters (Fig. 2).

4) *Sampler*: A simple D-flipflop or more advanced designs utilizing instance transmission gates can be used to sample CTBV signals at high speeds. This can be done to downconvert signals to lower rates, as mentioned before, or it may even find other uses purely in the CTBV domain. By downconverting the signal with samplers and doing the rest of the signal processing

in the system with regular digital logic, it is possible to create hybrid systems which take the best properties from both worlds.

C. Advantages

CTBV signal processing has several advantages over the other signal-processing domains, and we will now try to list and elaborate on them.

1) *Advanced Processing*: CTBV signal processing enables nonlinear processing operations which analog signal processing does not provide and which traditional digital signal processing only provides at much lower frequencies. We can perform simple pattern recognition in time and across several input signals; we have the delay line, which is a simple form of short-term memory, and we can easily conduct high-speed sampling to continue the signal processing using regular digital logic, which allows for interesting hybrid system designs.

2) *Simple*: The most basic CTBV circuits are simply logic gates or inverters, possibly with asymmetric NMOS/PMOS devices. This gives area-efficient implementations and low-power solutions, since such digital CMOS circuits only draw power when switching, unlike analog circuits which often have large bias currents flowing.

3) *High Speed*: Many CTBV operations will only be limited by the gate delays of the technology being used, and in modern CMOS technologies, this may only be a few tens of picoseconds. By using the difference of delays instead of absolute delays, we can even create operations which have relative time resolutions way below the gate delays.

4) *No Clock*: Since CTBV signal processing does not use clocking, except perhaps at the low frequency for hybrid systems, we save power and area for clock distribution. The operations will also be much faster since they do not need to wait for the next clock period to complete.

5) *Low Power*: CTBV signal processing is low power since, as previously mentioned, it uses no clock, is based on simple digital logic without bias currents, and has very simple and small circuits.

6) *Perfect for CMOS*: Since fine-pitch CMOS is cheap and well suited for high-speed digital circuits but not so good for analog designs, it is a suitable technology in which to implement CTBV systems. In the following text, we will show a measured CTBV sampler operating in excess of 30 GHz, which is a ten-fold increase compared to state-of-the-art clocked microprocessors.

D. Challenges

There are also a few challenges to keep in mind when working with CTBV technology.

1) *Limited Functionality*: CTBV signal processing will probably not be a general replacement for other signal-processing domains since its range of operations is limited and the signal coding is somewhat unusual.

2) *Minimum Pulse Width*: Pulses sent through a CTBV system must have a minimum width in order to avoid signal loss. For instance, two opposite transitions following each other closely may become even closer after going through a circuit since the first transition will leave the output a bit off

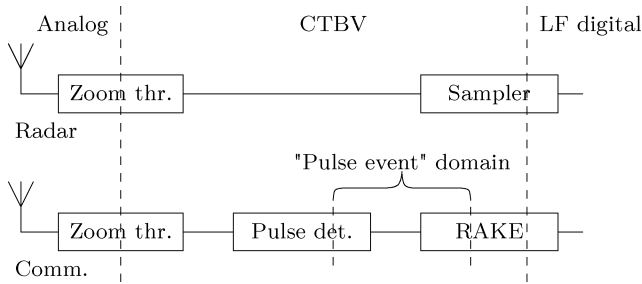


Fig. 3. UWB-IR signal-processing domains.

the rail, thus giving the second transition a head start on pulling the output to the other direction. This will cause the second transition to propagate faster through the gate than it should have. A second possible effect is the difference between rising and falling edge propagation times due to device mismatch. Due to these effects, we must make sure that the pulses in the input signals are sufficiently wide.

3) *Phase Noise*: The phase noise introduced by the gates will be the limiting factor for the SNR of a CTBV system. If the width of a pulse is considered to be the signal, the phase noise will be the noise when calculating the signal-to-noise ratio (SNR).

4) *Mismatch*: Device mismatch will cause incorrect delays in delay elements and can cause a difference in propagation delays for rising and falling edge transitions. CTBV systems may therefore be a demanding design effort.

E. CTBV for UWB-IR Applications

As we have indicated, CTBV is an interesting high-speed signal-processing domain with processing operations suitable for UWB-IR applications, such as simple pattern recognition for detecting UWB pulses. It allows for easy implementation of UWB-IR systems with multigigahertz signal processing in standard, cheap CMOS technology, enabling the production of affordable UWB-IR sensor and communication systems for biomedical and other applications.

In addition to being used for high-speed front-end processing in UWB-IR systems, CTBV is also probably suitable as a native representation for baseband UWB-IR signals. That is, after pulse detection in a UWB-IR receiver system, we are in a "pulse event domain" (Fig. 3), where the signal conveys "pulse events in time," which matches the CTBV signal representation very closely.

It is interesting to note that CTBV coding is actually also found in biology as coding in the nervous system (i.e., a nerve impulse is continuous in time, but has no information stored in its amplitude), and can thus be said to be a CTBV signal. Neuromorphic design [10] is performed in this type of domain.

In summary, CTBV is a rarely used, but is a very interesting signal-processing domain of its own, with a unique set of possibilities and limitations. It will probably not replace the more traditional signal-processing domains, in general, but it can be used to create simple and efficient hybrid systems together with these. In the following text, we will show how CTBV coding may be used for impulse radio technology.

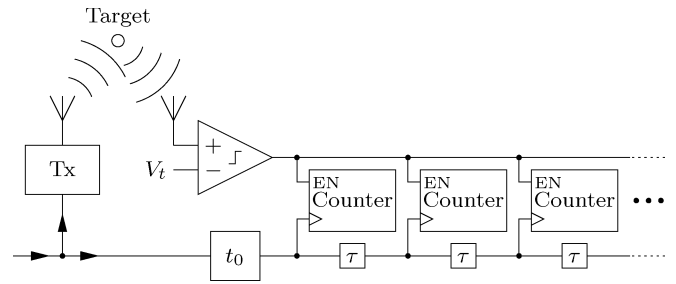


Fig. 4. Swept threshold sampler for radar.

V. SWEPT THRESHOLD RADAR

The swept threshold radar [8] works by transmitting UWB impulses, then thresholding and sampling the received echo signal. The digital but unclocked output from the thresholder is sampled in sequence by a number of parallel D-flip-flops, all having the thresholder output as their inputs. The D-flip-flops are set up to trigger one after another in rapid succession, with two inverters used as a delay element between them. When the signal has been sampled, a counter connected to each D-flip-flop is incremented if a "1" was sampled.

Another, more high-level way of looking at this architecture is to note the integration action of the latch-and-count structure. As shown in Fig. 4, these integrators are clocked at high speed, ensuring sampling at different depths with high resolution.

By repeating this procedure while sweeping the threshold level from below to above the range of the input signal, a full readout of the received signal is performed. The integrating action is achieved through repetition since the sampling is precisely timed. A counter which is clocked, for instance, at a high peak of the input signal, will sample many "1"s since the signal is above the threshold most of the time. A counter clocked at a low peak of the input signal, will sample few "1"s since the threshold quickly crosses the level of the low peak and the threshold output thus will be "0" for that low peak for most of the sweep. All in all, this architecture samples the received signal completely, which is assumed not to change too much during the sampling, at a high sampling rate given by the delay between the clocking of two adjacent counters. By removing the clutter of static objects, for instance, by high-pass filtering the entire readout-curve in postprocessing, we obtain a radar readout showing movements in front of the antennas.

An important detail is the integration of the white noise which is added to the received RF signal. This does mean that the counter values will be a little off from their expected values, and when repeating the sampling, the counter values will jump slightly up and down. By increasing the number of pulses sent during the sweep, the noise can be averaged down to an acceptable level. An ideal analog averaging sampler would have the noise reduced according to

$$\sigma_{N\text{recovered}} = \frac{\sigma_N}{\sqrt{n}} \quad (1)$$

where σ_N is the standard deviation of the noise in the input signal, n is the number of repeated samples, and $\sigma_{N\text{recovered}}$ is the standard deviation of the noise in the recovered signal (i.e., after averaging). We can see that by, for instance, increasing the

number of repeated samples by a factor of 100, the amplitude of the noise in the readout will be reduced by a factor of 10. Although this is the behavior of an ideal analog average sampler, the swept threshold sampler actually behaves quite similar, and for noisy signals, its performance is even quite close to the performance of the ideal analog average sampler [8].

The “digital” integrators (counters) are area efficient. Simultaneous sampling at a large number of different depths is easy to achieve by simply duplicating the integrator and sample with a delayed clock. The McEwan MIR only samples a single depth, which, of course, can be shifted back and forth (ranging) to obtain a full readout, but this procedure is time consuming. Since the sample rate is only determined by the delay between the triggering of the different counters, it can be very high, resulting in improved depth resolution. By using two inverters as a delay element in 90-nm CMOS, we have achieved an approximate 35-GHz sample rate, equivalent to a depth resolution of approximately 4.3 mm. Another important property of the digital integrator is the integration time. Unlike a lossy analog integrator, the “digital” integrator is lossless and when appropriate, the integration time may be arbitrarily long provided the counter is designed with a sufficient number of bits.

A. Zoom Thresholder

To convert the analog input RF signal to the CTBV domain, we use a thresholder (comparator) (i.e., a circuit which yields a “1” on its binary output when the analog input is above a given threshold V_t) and “0” otherwise. The threshold level should be tunable and the circuit should be able to react to small changes in the input signal at high frequencies.

We have already presented a CTBV thresholder in [8], and we will now introduce an improved design called the zoom thresholder. The circuit takes its name from the fact that the mechanism which sets the threshold level has a few distinct “zoom levels,” enabling a selectable tradeoff between the tuning range and precision of the threshold level.

The basic principle on which the zoom thresholder is based is shown in Fig. 5. A preamplifier amplifies the input signal by a small amount before it is coupled through a capacitor with a dc-setting resistor on the other side. The dc value of this node will be V'_t , and the amplified input signal will then swing around this voltage. Whenever the voltage of this node is above the operating point of the thresholding amplifier, its output will be “1,” and otherwise, it will be “0.” If, for example, the operating point of the thresholding amplifier is 0.5 V and V'_t is set to 0.4 V, the amplified input must reach a peak of at least 0.1 V in order to obtain a “1” on the output. And the higher the preamplifier gain is, the lower the peak in the actual input signal will have to be. A second effect of the coupling capacitor and the dc-setting resistor is that they will act as a high-pass filter for the input signal, thus rejecting low-frequency signals which could otherwise cause masking problems. By using an equivalent resistance circuit with a tunable value instead of a resistor, for example, a transconductance amplifier, it is possible to make the cutoff frequency of the high-pass filter tunable.

For the thresholding amplifier, we want a gain which is as constant as possible within our signal band. To implement this constant gain thresholding amplifier, we use a cascade of low

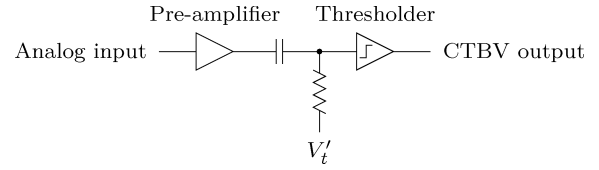


Fig. 5. Thresholder with an adjustable threshold level.

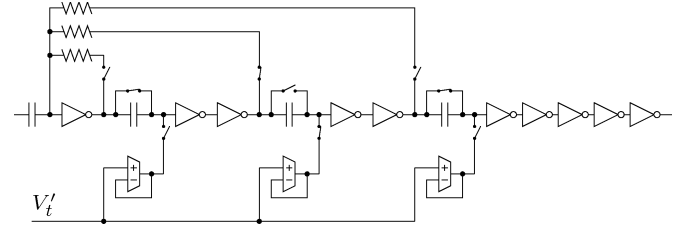


Fig. 6. Zoom thresholder.

gain common source elements designed in 90-nm CMOS, each with a gain of approximately 2.5. Ten such elements in cascade will provide an accumulated gain of approximately 10 000.

The primary reason for redesigning the thresholder presented in [8], was to reduce the flicker noise experienced with the previous design. The solution to this problem is quite simple—the signal must be amplified and configured with auto-zero circuits before the thresholding amplifier, as shown in Fig. 5. We see that the higher the preamplifier gain is, the lower the flicker noise introduced by the thresholding amplifier will be compared to the input signal. So by amplifying the input signal as much as possible without causing any other problems, the flicker noise will be reduced by a factor corresponding to the gain of the preamplifier. Simple auto-zero feedback keeps the signal within the operation range of the circuits. Errors in V'_t , such as noise and digital-to-analog converter (DAC) quantization errors, will also be reduced by the same factor. The ac coupling before the thresholding amplifier will filter out any flicker noise created by the preamplifier. Note that if the input signal is amplified too much, though, the tuning range of V'_t will not be able to cover the full extent of the input signal, and the circuit will thus not perform as expected.

To allow for selectable preamplifier gain, we have designed the zoom-thresholder circuit as shown in Fig. 6. The gain through the circuit is always ten cascaded common-source amplifier elements, but we can now select the number of amplifier elements to be used for preamplification, and then use the rest as a thresholding amplifier. The gain of this comparator may be too low if few amplifier elements are left for this stage. However, since the input to the thresholding amplifier then will be significantly amplified, the combined circuit should still have the gain expected from a thresholder. If V'_t in such a case should be set near the operating point of the amplifier elements, the output of the circuit should be somewhere between “1” and “0,” which would not be as expected. Due to thermal noise on the input though, the output would probably switch rapidly back and forth between the two, giving the binary output we want from the thresholder. The preamplifier will be self-biased with a feedback resistor which must be chosen with care to ensure stability of this stage. This feedback will make the

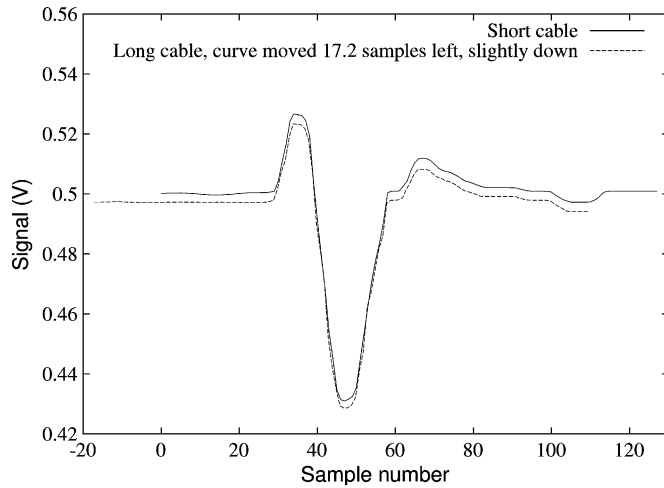


Fig. 7. Sample rate measurement.

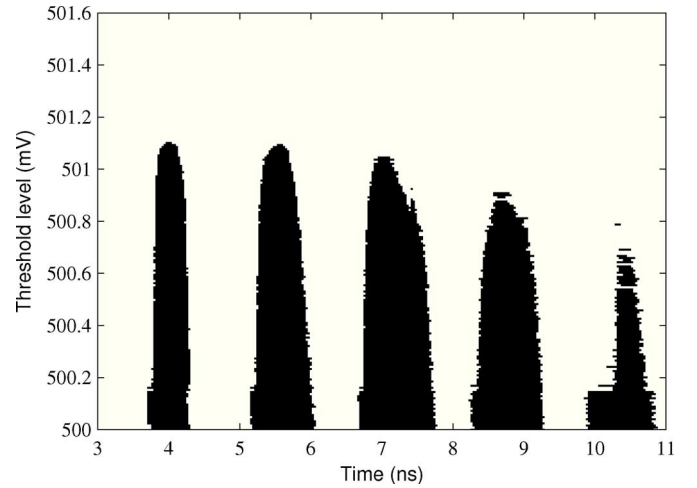


Fig. 8. Threshold sweep measurement.

preamplifier high pass and flicker noise strongly attenuated. The threshold setting resistor from Fig. 5 has been replaced by an equivalent transconductance amplifier, but with the input signal V_t' now isolated from the threshold circuit. To select the number of amplifier elements in the preamplifier stage, simple transmission gates are used to configure the circuit accordingly.

In summary, the zoom thresholder is a thresholder circuit with a tunable threshold level and a selectable “zoom factor” and combines low-noise amplifier (LNA) functions with comparator functions. The higher the zoom factor is, the smaller the threshold level tuning range will be, which means that a DAC with limited resolution controlling the threshold level will essentially get smaller steps, thus giving finer-grained control over the threshold level. In addition, a higher zoom factor will also reduce the flicker noise of the circuit more than a lower zoom factor will.

B. Measurements

The presented swept threshold radar principle has been successfully implemented in 90-nm CMOS, resulting in an approximate 1 mm² single-chip radar. The design uses the zoom thresholder which we have just described, and has 128 parallel samplers (i.e., counters, each with 24-b resolution). The samplers are triggered in sequence, each of them two inverter delays apart. We will now present some measurements of this implementation.

1) *Sample Rate Measurement:* Fig. 7 shows a simple readout from the radar with a loop cable and a 20-dB attenuator connecting the transmitter and the receiver. The lowest zoom level of the zoom thresholder has been used for these measurements. On the first axis of the plot, we have the sample numbers, 0 to 127, and on the second axis, we have the signal voltage. The actual voltage levels are only estimated values, so they might be a little off. An initial delay t_0 of 100 unit delays, each being two inverter delays long, has been used since the loop cable, about 1 m long, introduces a considerable delay to the signal. By performing the readout with two different loop cable lengths, we are able to determine the sample rate of the radar, since we will then know the difference in propagation

time for the signal and how many samples that difference corresponds to. The two different cables used were about 1 m long and had a length difference of 10.3 cm. The cable type was “Suhner Switzerland RG 400/U,” with a signal propagation velocity of 69% of c . As shown in the figure, the two readouts are almost identical when the curve of the long loop cable is moved 17.2 samples to the left. The difference in propagation delay through the two different cable lengths will be $\Delta t = 10.3 \text{ cm} / (0.69 \cdot 300\,000 \text{ km/s}) = 498 \text{ ps}$, which gives a sample rate of $f_s = \Delta \text{samples} / \Delta t = 17.2 / 498 \text{ ps} \approx 35 \text{ GHz}$. This corresponds to a depth resolution of $\Delta R = cT_s/2 = 300\,000 \text{ km/s} \cdot (1/35 \text{ GHz})/2 = 4.3 \text{ mm}$.

2) *Threshold Sweep Measurement:* Usually, the threshold is swept from a min to max value, and then the counter values are only read out after this sweep has been completed. By selecting one threshold level at a time, sending only one pulse and then reading out the counters, we are, however, able to see the details of the swept threshold sampling. Fig. 8 shows such a measurement. An antenna was connected to the radar, the radar was hung in free air, and the highest zoom level was used for these readouts. Several 128-sample windows have been stitched together to provide the 8-ns view shown in the plot. This is simply accomplished by repeating the readout with different initial delays t_0 . The threshold level was swept in discrete steps over a range covering the upper parts of the input signal. Each step corresponds to a horizontal line in the plot, each of which are essentially the CTBV output from the thresholder, sampled at $f_s = 35 \text{ GHz}$. Black means “1,” (i.e., the input signal is above the threshold and white means “0”), that is, the input signal is below the threshold. We see how the threshold sweeping scans out the shape of the received signal, and it is easy to see how a vertical integration of this plot, which is the normal mode of operation, will give a complete readout of the received signal. Note how the edges of the shapes are not completely smooth. This is due to noise, and can, as described earlier, be compensated for by sending several pulses per threshold level, thus averaging away the noise. Note also, again, that the voltage levels on the second axis are estimates, and may therefore be a little off from the real values.

3) *Noise Averaging Measurement:* Fig. 9 shows the effect of increased noise averaging (i.e., increased number of pulses sent

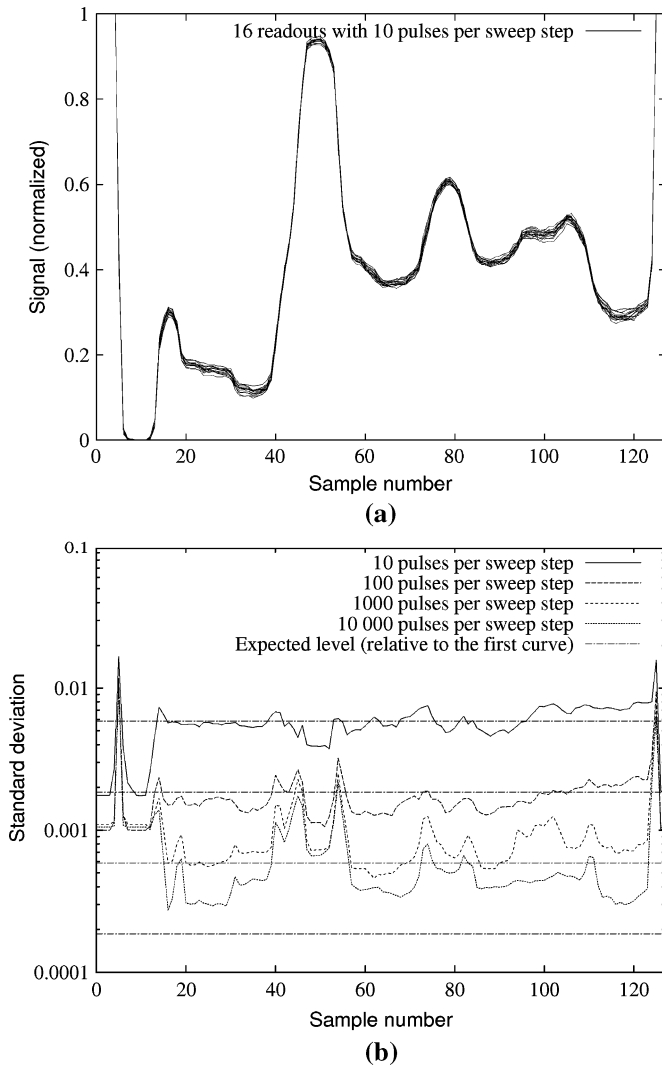


Fig. 9. Noise averaging measurement.

per sweep step). In Fig. 9(a), we see 16 repeated readouts, with only ten pulses per sweep step, plotted on top of each other. This gives a blurry curve, illustrating the noise in the readout. The more pulses per sweep step, the less blurry the curve will be. In Fig. 9(b), the standard deviation for each of the 128 samples has been calculated from 256 separate readouts. Simply put, this is a measure of how blurry the first curve is, calculated for each point along the curve. Results from four different pulses-per-step settings are shown.

We see that the more pulses we send per sweep step, the less the standard deviation becomes, thus confirming that this actually reduces the noise. Some horizontal lines are plotted in Fig. 9(b) which indicate the expected standard deviation levels of the four different pulses-per-step settings. These lines are simply $1/\sqrt{n}$ -lines relative to the upper curve. Although an increased number of pulses per step reduces the noise, the reduction is not quite as large as expected for the highest averaging settings. This is perhaps due to remaining flicker noise in the zoom threshold, which would create some low-frequency noise components which cannot be averaged away by the rapidly successive samplings we are using. To compensate this to some extent, we

have, previous to the standard deviation calculation, removed the dc component of the curve. A limitation of this approach though is that the threshold sweeping takes a given amount of time, during which the value of the low-frequency noise may have changed slightly. This can cause the lower and upper parts of an input signal to obtain different “dc-noise” values, thus rendering the dc-removal technique less effective.

We can also see that steep slopes in the input signal cause an increase in the standard deviation. This is probably because of jitter noise, since even a slight time shift of the signal would cause a large amplitude shift at the slopes.

4) *Distance Measurement:* We are working to establish accurate distance measurements using the impulse radar. Crude manual testing indicated distance measurements with millimeter resolution are feasible with RF. The main reason is the high-speed samplers tracing out the received signal in time.

C. Biomedical Sensor Applications

The short-range impulse radar is basically a novel close range EM sensor with high depth resolution. Vital signs, such as heartbeats and breathing, have already been shown to work. Unlike alternative vital sign sensors (ultrasonic or electric), the impulse radar does not have to be directly attached to the body usually appreciated by the users. A number of interesting applications are appearing, such as the sleep warning in cars or an “intelligent mattress” for hospital beds. Penetration properties of impulses may facilitate accurate distance measurements to heavy matters and even landmine detectors. The miniature size of 1 mm² silicon operating at a 1-V power supply enables a number of compact, battery-operated gadgets.

VI. IMPULSE RADIO COMMUNICATION

It is also possible to explore impulse radio for communication. Significant efforts have been invested in the IEEE 802.15.4a standard for low-rate wireless personal-area networks (WPANs) [11]. One of the proposed physical layers is impulse radio based, using continuous pulse ultrawideband (C-UWB). In the literature, a number of coding schemes are proposed [12] and evaluated. In this paper, we will not treat or evaluate different coding techniques although they are of uttermost importance, but concentrate on how the CTBV paradigm may be explored for power efficient symbol recovery in a correlating RAKE receiver.

For the purpose of understanding the CTBV processing in the receiver, we assume a simplified coding scheme where we have a small number of symbols such as “0” and “1”. Each symbol (PN_0 and PN_1) is coded as a sequence of subnanosecond pulses [12], [13]. The desired noise-like property is achieved by distributing the pulses in a pseudorandom pattern in time. The actual pulse patterns for each symbol are often constructed as a clocked sequence of “chips” (time hopping) with actual emissions in those chips where PN bits are set. By using pulses, such as higher order Gaussians, proper filling of the Federal Communications Commission (FCC)-approved frequency mask is possible. The coding scheme and simplified computational elements are shown in Fig. 10.

From an implementation perspective, the transmission of symbols as PN sequence is simple, requiring no oscillators

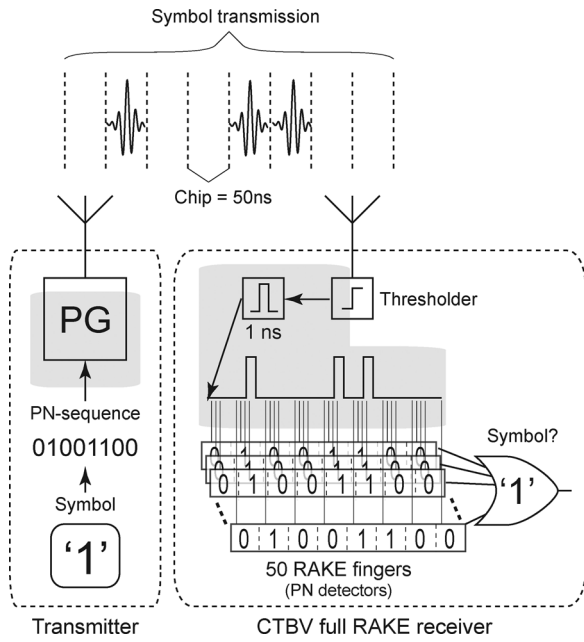


Fig. 10. Simplified view of the communication scheme, using an 8-b-long PN code as an example. The shaded parts utilize CTBV coding.

for carrier generation. These carrier-free transmitters can be made to be very power efficient and were provided by the company Novelda AS [14]. Their indicated power consumption is $< 500 \mu\text{W}$ at a PRF of 20 MHz. Unfortunately, it is hard to make a power-efficient receiver due to the statistical symbol recovery required (i.e., the RAKE), and due to the power consumption of the input stage.

One major issue is synchronization. Incorporating the clock in the pseudonoise (PN) sequence will most certainly give undesirable spectral peaks. Aiming for UWB-IR solutions, sampling rates at several gigahertz are hard to make and are very power demanding. Again, by exploring the concepts of CTBV-coded signals, we are able to find a power-efficient architecture to avoid precise clock synchronization.

The concept of RAKE architecture was proposed a long time ago [15] and sifts through the incoming signal for PN codes looking for line-of-sight (LOS) and nonline-of-sight (NLOS) PN patterns. This pure time-domain processing structure is equivalent to running cross-correlation with some expected template. With short pulses, patterns due to multipath transmission may also be recovered and used constructively for symbol detection. The LOS PN and the NLOS PNs are usually detected in RAKE fingers. The RAKE function is often implemented as software on DSPs with typically 2–3 RAKE fingers, known as partial RAKE receivers.

To facilitate understanding of how CTBV signal processing can be utilized for power-efficient RAKE receiver implementations, the receiver signal flow is shown as follows.

- 1) The incoming RF signal is filtered and immediately thresholded (comparator) by using a tunable threshold level. This conversion (1-b ADC) is continuous in time and encodes the incoming signal into the CTBV domain.
- 2) Since the actual RF pulses are higher order Gaussian pulses consisting of a few cycles, we “merge” the suc-

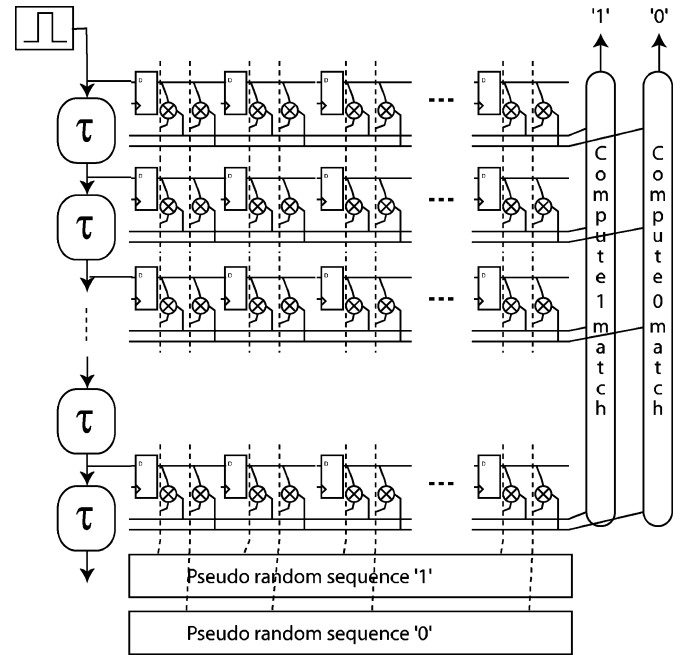


Fig. 11. Thresholding, continuous-time full RAKE receiver.

cessive peaks together to a 1-ns CTBV pulse. Each transmitted higher order Gaussian pulse will now appear as a 1-ns CTBV pulse and the CTBV sequence should match the transmitted PN sequence.

- 3) The CTBV sequence of pulses is then “stored” in a delay line matching the chip length.
- 4) With taps of the delay line for every nanosecond (Fig. 11), we may sample the taps with a chip-rate clock, thus recovering the expected PN sequence.

The CTBV thresholding RAKE receiver architecture [16] is shown in Fig. 11. The efficiency of CTBV circuits enables implementation of a “full” RAKE receiver, alleviating rigorous synchronization. Sampling the delay line with the chip-rate clock (typically in the megahertz range), we are able to latch the signal value as a bit in a clocked shift register. The LOS PN pattern may occur in any of the RAKE fingers constituted by the shift registers. In the end, we really do not care which one, since we are sampling them all. We may even look for our desired PN pattern of the NLOS signals by considering matching in all RAKE fingers. Correlating bit patterns is also simple. The multiplications of cross correlation are reduced to simple AND gates. The clock accuracy is only required for the duration of one PN-coded symbol, acceptable for most simple crystal oscillators.

So by using thresholding and continuous time signals, we are able to design a power-efficient impulse radio, correlating the receiver to exploring a full RAKE architecture.

The impulse radio transmission scheme has a number of interesting features as follows.

- 1) Improved sensitivity. One of the major improvements for robust communication is increased sensitivity. The “integration” action of repeated pulse transmission of the swept threshold radar is equivalent to increasing the number of pulses of the PN codes by increasing the length of the PN

sequence. Intuitively, by having more pulses to look for, we achieve a higher level of signal integration. Evidently, longer PN sequences will reduce the transmission bandwidth, but we are gaining sensitivity. To what extent this technique will outperform the sensitivity of carrier-based receivers remains to be seen, but the potential is certainly there.

- 2) Channelization. It is possible to tune UWB-IR transmission for a large number of channels (simultaneous transmissions). If sparsely populated PN sequences are used, a large number of orthogonal PN sequences are available, enabling a large number of channels (only two symbols are required for each channel). Getting more than 100 channels is easy and with minor interference. Compared to typically 16 channels in a narrowband systems, UWB-IR is adapting to densely populated environments (such as hospitals).
- 3) Multipath exploitation. In contrast to carrier-based systems, UWB-IR systems generally benefit from multipath transmissions due to the short pulse duration (subnanoseconds) [13]. This allows us to reduce fade margins [17]. The full RAKE structure enables the combination of most multipath components. In fact, reflections are improving channel quality.
- 4) Power efficiency. Exploring the CTBV domain for UWB-IR-based transceivers facilitates an extremely low-power transmitter and also a correlating receiver for power-efficient implementation in CMOS. The CTBV delay line solves the synchronization problem as well as high-speed clocks for gigahertz sampling.

A single-chip, fully functional UWB-IR transceiver in CMOS is, to our best knowledge, still not available, although several companies have worked toward these solutions for a long time. We believe novel processing paradigms, such as CTBV signal processing, are required for nanometer technology. The startup company Novelda AS [14] is pursuing functional CMOS transceivers based on the CTBV paradigm.

VII. UWB-IR RANGING TRANSCEIVER

The functionality of the presented impulse radar may be combined with transceiver technology. As indicated, millimeter distance measurements are achievable by using the high-speed sampler. The concept of "active echo" was developed by Andersen [18] based on the "passive echo" of the impulse radar. In a master-slave configuration, the slave is actively echoing a transmitted pulse from the master node. The master node is using the high-speed sampler for accurate time-of-flight measurement. Knowing the constant delay added by the slave, the master node can compute the actual time of flight and thereby the distance between the nodes with high accuracy. Configuring these ranging transceivers in a wireless sensor network facilitates localization, provided that a sufficient number of nodes are within range [19]. We expect these ranging UWB-IR transceivers to outperform current RSSI-based distance measurement with two to three orders of magnitude.

VIII. CONCLUSION

The potential for UWB-IR in biomedical communication and sensing is certainly promising. In this paper, we have

shown the fundamental opportunities of UWB-IR and how technology-friendly solutions exist. The novel CTBV design paradigm has been proven to work for impulse radar implementations. Power-efficient correlating receiver solutions are also proposed. The accurate distance measurements of the impulse radar may be combined with a UWB-IR transceiver by adding accurate ranging capabilities to wireless technology. There are strong indications that UWB-IR technology has superior wireless functionality for biomedical applications.

ACKNOWLEDGMENT

The authors would like to thank the startup company, Novelda AS [14], for their generous support in putting together a working impulse radar system used in our measurements. The technical staff has contributed through a number of valuable discussions and suggestions.

REFERENCES

- [1] B. Warneke, M. Last, B. Liebowitz, and K. Pister, "Smart dust: Communicating with a cubic-millimeter computer," *Computer*, vol. 34, no. 1, pp. 44–51, Jan. 2001.
- [2] S. Gezici, "Theoretical limits for estimation of periodic movements in pulse-based uwb systems," *IEEE J. Sel. Topics Signal Process.*, vol. 1, no. 3, pp. 405–417, Oct. 2007.
- [3] S. Venkatesh, C. Anderson, N. Rivera, and R. Buehrer, "Implementation and analysis of respiration-rate estimation using impulse-based uwb," in *Proc. IEEE Military Communications Conf.*, 2005, vol. 5, pp. 3314–3320.
- [4] N. V. Rivera, S. Venkatesh, C. Anderson, and R. M. Buehrer, "Multi-target estimation of heart and respiration rates using ultra wideband sensors," presented at the XIV Eur. Signal Processing Conf., Florence, Italy, Sep. 4–8, 2006. [Online]. Available: <http://www.eurasip.org/Proceedings/Eusipco/Eusipco2006/papers/1568987536.pdf>.
- [5] T. E. McEwan, "Ultra-Wideband radar motion sensor," U.S. Patent 5 361 070, Nov. 1, 1994.
- [6] S. Azevedo and T. E. McEwan, "Micropower impulse radar," *IEEE Potentials*, no. 2, pp. 15–20, Apr./May 1997.
- [7] E. M. Staderini, "Uwb radars in medicine," *IEEE Aerosp. Electron. Syst. Mag.*, vol. 17, no. 1, pp. 13–18, Jan. 2002.
- [8] H. A. Hjortland, D. T. Wisland, T. S. Lande, C. Limbodal, and K. Meisal, "Thresholded samplers for uwb impulse radar," in *Proc. IEEE Int. Symp. Circuits and Systems*, 2007, pp. 1210–1213.
- [9] T. S. Lande and H. A. Hjortland, "Impulse radio technology for biomedical applications," in *Proc. IEEE Biomed. Circuits Syst. Conf.*, 2007, pp. 67–70.
- [10] C. Mead, *Analog VLSI and Neural Systems*. Reading, MA: Addison-Wesley, 1989.
- [11] [Online]. Available: <http://www.ieee802.org/15/pub/TG4a.html>.
- [12] M.-G. D. Benedetto and G. Giancola, *Understanding Ultra Wide Band Radio Fundamentals*. Upper Saddle River, NJ: Prentice-Hall, 2004.
- [13] M. Win and R. Scholtz, "Impulse radio: How it works," *IEEE Commun. Lett.*, vol. 2, no. 2, pp. 36–38, Feb. 1998.
- [14] [Online]. Available: <http://www.novelda.no/>.
- [15] R. Price and P. Green, "A communication technique for multipath channels," in *Proc. IRE*, 1958, pp. 555–570.
- [16] C. Limbodal, K. Meisal, T. S. Lande, and D. Wisland, "A spatial rake-receiver for real-time uwb-ir applications," in *Proc. IEEE Int. Conf. Ultra Wideband*, 2005, Cat. no. 05EX1170C.
- [17] F. Ramirez-Mireles and R. Scholtz, "System performance analysis of impulse radio modulation," in *Proc. IEEE Radio and Wireless Conf.*, 1998, pp. 67–70.
- [18] N. Andersen, "Active echo, high precision ranging in wireless sensor networks," Master's dissertation, Dept. Inf., Univ. Oslo, Oslo, Norway, 2007.
- [19] H. K. Olafsen, "Wireless sensor network localization strategies," Master's dissertation, Dept. Inf., University of Oslo, Oslo, Norway, 2007.



Håkon A. Hjortland received the M.Inf. degree in microelectronics from the University of Oslo, Oslo, Norway, in 2006, where he is currently pursuing the Ph.D. degree.

His research interests include UWB impulse radio, short-range radar, and wireless data communication.



Tor Sverre "Bassen" Lande (M'93–SM'06) is currently a Professor in the Microelectronic at Department of Informatics, University of Oslo, Oslo, Norway. In 2004 he was appointed Visiting Professor at the Institute of Biomedical Engineering, Imperial College London, London, U.K. His primary research is related to microelectronics, both digital and analog. His research fields are neuromorphic engineering, analog signal processing, micropower circuit design, biomedical circuits and systems, and impulse radio. The understanding of representation

and computation in different computational paradigms has spawned novel ways of designing both large and smaller systems using state-variables like frequency modulation and/or spike coding. Driven by the demands from practical application focus on technology, especially micropower circuit biasing and floating-gate structures in combination with signal processing techniques like log-domain, has been focused. With the Imperial College relation more emphasis has been on microelectronics for implantable and wearable systems. Recently, short range impulse radio for low-power personal area networking and high resolution radar in biomedical applications is under investigation. He is the author or coauthor of more than 90 scientific publications with chapters in two books. He is currently serving as an Associate Editor of several scientific journals.

Dr. Lande has served as a Guest Editor of several journals, including *Analog Integrated Circuits and Signal Processing*, where he is also serving as a member of the Editorial Board. He is a Technical Committee Member of several international conferences and has served as reviewer for a number of international technical journals. He has served as Technical Program Chair for several conferences (ISCAS 2003 in Bangkok, Thailand, NORCHIP 2004 in Oslo, Norway, BioCAS Workshop 2004 in Singapore, BioCAS 2006 in London, U.K.). He was chair elect (2003–2005) of the IEEE Biomedical Circuits and Systems technical committee (BioCAS) and is also a member of other IEEE Circuits and Systems Society technical committees. In 2006, he was appointed Distinguished Lecturer of the IEEE Circuits and Systems Society (CAS-S) and elected member of CAS Board of Governors. He is the current Editor-in-Chief of the new publication IEEE TRANSACTIONS ON BIOMEDICAL CIRCUITS AND SYSTEMS.

Paper 5 CMOS nanoscale impulse radar utilized in 2-dimensional ISAR imaging system

D. T. Wisland, S. Stoa, N. Andersen, K. Granhaug, T.-S. Lande, and H. A. Hjortland. “CMOS nanoscale impulse radar utilized in 2-dimensional ISAR imaging system”. In: *Radar Conference (RADAR), 2012 IEEE*, pp. 0714–0719, Atlanta, GA, May 2012. doi : 10.1109/RADAR.2012.6212231. © 2012 IEEE. Reprinted with permission.

My Contributions to This Paper

I was part of the team that developed the radar system which is presented and used for ISAR imaging in the paper. The radar is based on the swept threshold sampler principle which we present in other papers. I also helped review the paper. I was not involved with the ISAR imaging.

CMOS Nanoscale Impulse Radar Utilized in 2-Dimensional ISAR Imaging System

Dag T. Wisland, Stig Støa, Nikolaj Andersen,
and Kristian Granhaug
Novelda AS
Kviteseid, Norway
dag@novelda.no

Tor Sverre Lande and Håkon A. Hjortland
Dept. of Informatics
University of Oslo
Oslo, Norway
bassen@ifi.uio.no

Abstract—This paper presents a complete impulse based radar transceiver integrated on a single chip. The high bandwidth of the transmitted pulses offers unique penetration abilities and very high accuracy. By taking advantage of the Continuous Time Binary Valued (CTBV) design paradigm, some of the main show stoppers of systems based on traditional digital design techniques have been removed. This paper explores the advantages and possibilities of the fully integrated CTBV radar, and introduces basic background on the Inverse Synthetic Aperture Radar (ISAR) along with experimental results. Measurements have been performed that confirm the theory and show that the proposed nanoscale impulse radar system may be utilized for 2D-imaging combining low cost implementation with high quality. The CMOS radar consumes 113mW from a 1.2V/2.5V power supply.

Keywords - impulse radar, UWB, CMOS, ISAR, CTBV.

I. INTRODUCTION

The concept of impulse based radio transmission is the oldest communication technology explored by Marconi in 1901 using wideband sparks for wireless communication. The pulse based communication was however soon replaced with narrowband, carrier-based solutions still dominating present day communication and radar systems. However, the U.S. regulating authority (FCC) has released the 3.1-10.6 GHz frequencies for unlicensed use, provided that emission levels are kept low (-41.3 dBm/MHz). This new unlicensed band, often called ultrawideband (UWB), is the widest unlicensed frequency band ever released, and is wide enough for proper impulse radio applications. Some reported results of using the impulse-based systems may be found in [1]–[3]. One noticeable exploration of impulse radio technology is the work by Tom McEwan on Micropower Impulse Radar (MIR) [4], [5], where small energy bursts are emitted as impulses (noise) and the very weak reflected energy is recovered using range gating in combination with significant analog integration. Staderini [3] and [6], was also able to remotely recover heart beats by exploring the MIR architecture. Two important observations of these works are: 1) the penetration of heavy materials with pulses and 2) the ability to recover very small amounts of reflected energies using integration. Both of these features are also desirable from a communication perspective.

A major issue of the MIR technique is its fundamental analog nature in combination with the required high-speed sampling. The implementation of MIR in low-voltage CMOS technology is therefore not straightforward and alternative methods must be explored.

Generally, the main challenge of implementing an impulse radar system in CMOS is related to the extremely high speed requirements of the front-end components, and in particular the sampler. The main task of the sampler is to capture an analog input signal and store it as digitally readable values, and its design depends on requirements such as speed, power and resolution. In our case, the sampler had to capture a signal with a bandwidth of several gigahertz, which is a very demanding requirement to meet. By conventional methods, meeting this requirement would imply the running of high speed clocks, and consuming large amounts of power and silicon.

Rather than continuously sampling the received signal, the proposed topology employs a concept known as strobed sampling. In Fig. 1 the principle of an Impulse Radar using strobed sampling is shown. For each pulse transmitted, the backscattered EM energy is sampled after a given time offset. This offset represents the time-of-flight of the signal relative to the time of transmission, which in turn can be used to represent the distance to the remote object. In strobed sampling radar systems, highly accurate time offsets, as well as high-speed sampling, is required. These functions can be hard to realize with conventional synchronous digital methodologies. The core circuitry of the strobed sampler used in the proposed impulse radar therefore operates in the continuous-time domain, without the use of a clock. As shown in Fig. 1, this technique moves the point where the signal is converted from continuous to discrete time compared to conventional systems, where the sampling and time offset relies on a clock signal. This methodology is the key to achieving time-of-flight measurement accuracy in the picoseconds range without high frequency clock signals. Section II presents the CTBV implementation approach and Section III the theoretical background of the 2-D ISAR case study. Measured results are provided in Section IV, while the paper is concluded in Section V.

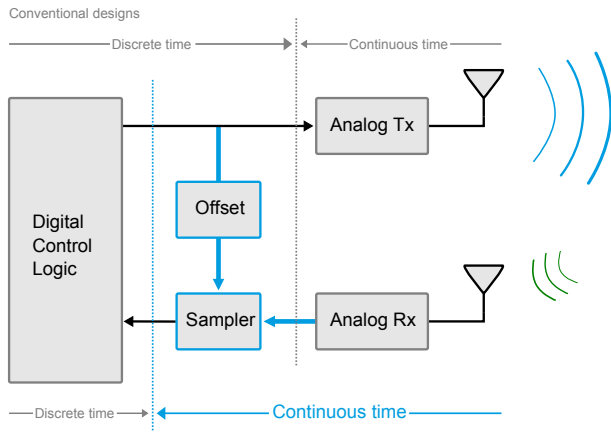


Fig. 1. Conceptual block diagram of impulse radar illustrating the principle of continuous-time strobed sampling compared to conventional sampling.

II. THE CTBV RADAR IMPLEMENTATION APPROACH

The strobed sampling technique briefly presented in the introduction does however only provide a single point in time. A trace of the received signal over time is required in order to correlate the received signal to a template when, for instance, searching for reflecting objects in the signal path. The proposed radar topology extends the basic strobed sampler by exploring the domain of CTBV coding. Unlike the regular binary valued signals that relate to a clock edge, the CTBV signal has near infinite resolution in the time domain. The analog to CTBV conversion is realized by a high speed 1-bit continuous-time quantizer.

In Fig. 2 the block diagram of the proposed radar system is illustrated. The transmitter is triggered by the digital control logic and transmits a 1st order Gaussian pulse. After a given time offset the high speed sampler is triggered. A 1-bit quantizer converts the incoming signal to a binary sequence. This CTBV coded signal is then distributed to a number of parallel samplers. As in the strobed sampler, the trigger signal is delayed with a time offset. This trigger signal is then connected to the parallel samplers, each with a slightly different time offset. This way, the samplers sample the incoming CTBV signal in rapid sequence. The output of these CTBV samplers can then be read as normal digital values.

A. The Average Sampling Technique

With this method the incoming signal will be sampled at a number of different depths (typically 512) for each transmitted pulse. Consequently, the proposed impulse radar is operating like a number of traditional radars in parallel. However, the previously described CTBV sampler only samples single bits due to the crude single-bit quantizer. In order to increase the resolution of the sampled signal, the signal is converted using a technique called swept threshold sampling [7]. This technique shares similarities with what is

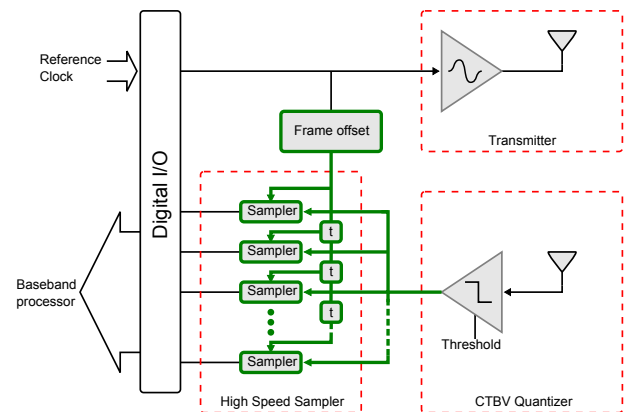


Fig. 2. Block diagram of the proposed nanoscale impulse radar illustrating the CTBV signal flow in green.

known as a flash or direct conversion ADC.

However, in contrast to the flash converter, only a single quantizer is used in the swept threshold converter. Instead, the sampling procedure is repeated while the threshold is swept over the area of interest. For each threshold step the conversion result is accumulated in a digital counter. In a noiseless environment, the resolution of this system is given by the step size of the threshold sweep. After the sweep is complete, the converted value can be read from the counter. Say, for instance, that the threshold was swept from 0 to 1V in steps of 10mV. For an input of 1V, the result of this conversion is 100. For an input of 0.6V, the result will be 60 and for an input of 0V the result is 0. For noisy or weak signals this procedure can be repeated several times, which adds processing gain in the form of averaging. Substantial averaging may be required to recover the received signal. Fortunately, the digital averagers are simply digital counters. In comparable systems, lossy analog integrators are used, limiting integration time. The digital counter-averager may be extended to count for a long time, only limited by the number of bits in the counters. In this way the proposed Nanoscale Impulse Radar opens up the possibility of adding a large processing gain early in the processing chain.

In addition to the above presented critical front-end components, the system comprises a baseband processor controlling range-gating and the threshold sweeping operation. The circuit also includes an SPI-controller to interface with a microcontroller or a PC via a USB interface. The chip is 2 x 2 mm in size and has been implemented in a standard 90 nm CMOS technology from TSMC.

III. 2-D ISAR IMAGING APPLICATION CASE STUDY

To evaluate the proposed impulse radar system, an evaluation board combining the CMOS radar IC with antennas and necessary support components was developed. The evaluation board as illustrated in Fig. 3 is connected to a computer

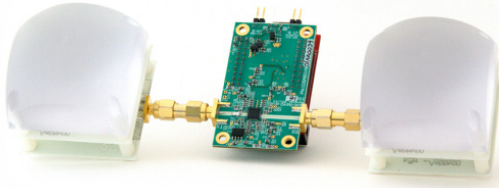


Fig. 3. Novelda NVA-R640 Evaluation Board used for case studies and system characterization. Antennas were replaced by SAS-571 for measurements.

through a USB interface and controlled using Matlab running the ISAR signal processing routines. The basic principle of ISAR imaging is that the impulse radar emits a very short pulse in time, which is reflected from different objects and received at the receiver antenna. Knowing the velocity of the wave and the time it takes for the reflection to reach the receiver, the distance the wave has traveled can be calculated using the well-known formula

$$D = v * t, \quad (1)$$

where v is the velocity of the wave, and t is the time the wave uses to hit the receiver. The total distance the wave travels is from the transmitter to the object, then from the object and to the receiver. The distance to the object is $D/2$ if the distance between the two antennas is small compared to the distance to the object, i.e. $D_{(TX \rightarrow RX)} \ll D_{OBJECT}$.

ISAR or Inverse Synthetic Aperture Radar is a technique used to generate a high resolution two dimensional image of an object using impulse radar. The image represents a cross section of the object under test. The ISAR technique utilizes the rotating movement of an object, and the resultant reflections, to calculate the two dimensional image of the object rather than moving the radar around the object as with SAR (Synthetic Aperture Radar). The name *inverse* comes from the fact that ISAR is the inverse of the SAR technique, and the name *synthetic* comes from the fact that the images are assembled from subsidiary elements [8]. To create the two dimensional image of the object under test, the reflection from each point for all angles needs to be evaluated and compared. This means that the distances to a certain point must be known for all angles. A reference point also has to be established to get exact information. An accurate measurement of the components in the set-up is important to ensure best possible image processing results.

A. Distance Calculations

Figure 4 shows that the object under test is represented here as a point P on a rotating platform, with the coordinates (x_p, y_p) . This point P will change coordinates as the platform rotates, and therefore the distance the wave travels will also

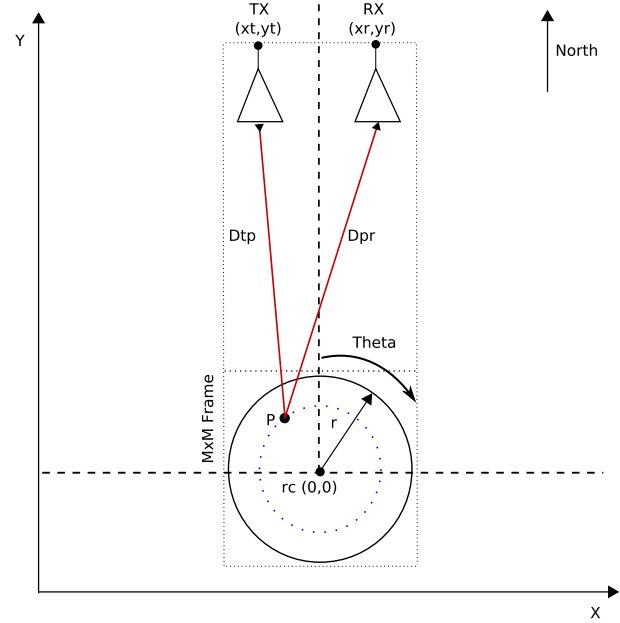


Fig. 4. Coordinate system setup for ISAR measurement.

change for different angles. The total distance the wave travels is simply the sum of the distance between the transmitting antenna and the object (D_{tp}), and the distance between the object and the receiving antenna (D_{pr}) [9]:

$$D = D_{tp} + D_{pr} \quad (2)$$

For a motionless situation the distances D_{tp} and D_{pr} can be calculated using the Pythagorean Theorem, resulting in the following expressions:

$$D_{tp} = \sqrt{(x_p - x_t)^2 + (y_p - y_t)^2} \quad (3)$$

$$D_{pr} = \sqrt{(x_p - x_r)^2 + (y_p - y_r)^2} \quad (4)$$

Knowing the initial coordinates of the point P it is possible to calculate the position for this point for all angles in the rotation. When the platform rotates, the coordinates to the point P is transformed into a function of the angle theta:

$$x_p(\theta) = x_p * \cos(\theta) - y_p * \sin(\theta) \quad (5)$$

$$y_p(\theta) = x_p * \sin(\theta) + y_p * \cos(\theta) \quad (6)$$

By substituting equation (5) and (6) into equation (3) and (4), these equations can, in combination with (2), fully describe the distance the impulse emitted from the antenna has to travel before hitting the receiving antenna for all angles of rotation. An important thing to remember, as shown in Fig. 4, is that the origin of the coordinate system is the center of rotation (rc), and the positive x-axis and positive y-axis is in the east and north direction respectively. If these coordinates are not measured correctly related to the frame, the produced

image from the DSP software will not correspond with the cross section of the object under test.

In order to calculate relative movement, we define the distance to the center of rotation, D_0 , as a reference point. Using (2)-(6), relative movement can be approximated by

$$\Delta D = D - D_0 = (\sqrt{(x_p(\theta) - x_t)^2 + (y_p(\theta) - y_t)^2} + \sqrt{(x_p(\theta) - x_r)^2 + (y_p(\theta) - y_r)^2} - (\sqrt{x_t^2 + y_t^2} + \sqrt{x_r^2 + y_r^2})) \quad (7)$$

Since the time delay and the distance are known, the corresponding phase shifts, ϕ , for frequencies, f , in the spectrum of the reflected pulses can be expressed as

$$\Delta\phi = 2\pi f\tau = 2\pi f \frac{\Delta D}{c} \quad (8)$$

In the time domain the signals carry information about the intensity of the reflected wave within the time window and the distance to these points. The time window is the window of time where the radar records reflected waves. The proposed Nanoscale Impulse Radar takes 512 samples within the time window with approximately 27 ps time difference, which is approximately 4 mm spacing between each sample, resulting in a time window of 2.048 meters. The resulting matrix after one revolution is then a $512 \times M$ matrix, where M is the number of steps. To get a quadratic matrix for the signal processing, and remove samples that are not covering the object under test, the excess data of the dataset is removed and results in the $M \times M$ frame shown in Fig. 4. To ensure that this frame covers the rotating platform, the length of the time window must be equal to or exceed the diameter of the rotating platform. This length can be verified when knowing the number of steps, which were 288 in this set-up:

$$M * 0.004m = 288 * 0.004m = 1.152m$$

The diameter (d) of the rotating platform used in this set-up was 600 mm, so the frame covered almost two times the diameter of the platform. The minimum number of steps needed to cover the 600 mm platform was 150 steps. The DSP software also includes two filters to remove noise from the frame in order to enhance image quality. The first filter is a DC, or low frequency, noise filter that filters out the low frequency noise of the frame by setting a zero mean in each frame. The second filter is a clutter removal filter that subtracts the static clutter map from the image to remove reflections from static objects in the surroundings, meaning that the static clutter map is a dataset of the environment without the object under test present. There are two ways of doing static clutter map removal:

- Single clutter map subtraction
- Multiple clutter map subtraction

Single clutter map subtraction uses only one initial static clutter map that is subtracted from all frames. This is not an ideal method since the intensity of static reflections change slightly as the platform rotates. Therefore, a multiple clutter

map subtraction can be used to further enhance the image quality. To utilize this method, a static clutter map must be captured for all angles, meaning a full revolution providing an $M \times M$ matrix with M static clutter maps that are subtracted from the respective frames. As a result, two complete revolutions will occur for each recording. One for creating static clutter maps for all angles, without the object present, and the other for measurement data of the object under test.

B. Pixel Intensity

To calculate the pixel intensity for every pixel (x_{pix}, y_{pix}) in the $M \times M$ matrix, an FFT is performed along the time variable dimension of the recorded data:

$$S_{f_n}^{\theta_k} = FFT\{s_{t_n}^{\theta_k}\} \quad (9)$$

where S denotes the Fourier transformed matrix, s the measurement matrix and θ_k, f_n and t_k are indexes of angle, frequency and time of flight respectively. Each sample of the transformed matrix S_f^{θ} is then multiplied with the corresponding phase shift derived from (8). This is to shift the reflected pulses from pixel coordinates (x_{pix}, y_{pix}) on the rotating platform for all angles back to the reference point:

$$\hat{S}_{f_n}^{\theta_k}(x_{pix}, y_{pix}) = S_{f_n}^{\theta_k} e^{-j\Delta\phi} \quad (10)$$

Finally the intensity of each pixel can be calculated from the inverse Fourier transform of \hat{S}, \hat{s} , by summing up the samples that correspond to the reference point $M/2$:

$$I(x_{pix}, y_{pix}) = \sum_k \hat{s}_{M/2}^{\theta_k} \quad (11)$$

IV. EXPERIMENTAL RESULTS

Several objects were tested grading from low complexity to higher complexity. Different initial coordinates were also used to confirm the usability of the DSP software. The dataset can be down-sampled with a factor X to reduce the runtime equally, but the trade-off is lower image quality. The images presented in these results are not down-sampled. It is to be noted that the reference line has been shifted around $M/2$ to get better results. This is due to inaccuracies in the measurement set-up. The reference line number is displayed at the top of the processed image.

The first object under test was a metal can with dimensions 115x60x200 mm placed at the center (0,0) of the rotating platform, see Fig. 5. The processed image shown in Fig. 6 clearly verifies these dimensions and coordinates. It is to be noted that there are some visible reflections from the rotating platform and the environments in the image, which could be reduced by addressing some of the sources of error discussed later in this paper.

The second object under test was a wooden box with dimensions 350x350x290 mm with an aluminum foil covered cardboard box inside. The box inside was rotated 45° relative to the wooden box, and the wooden box was placed at the

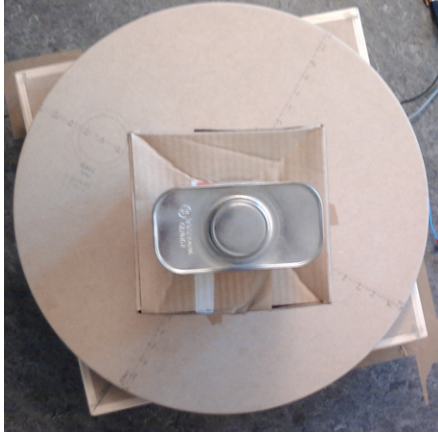


Fig. 5. Test setup 1 - Metal can on rotating platform.

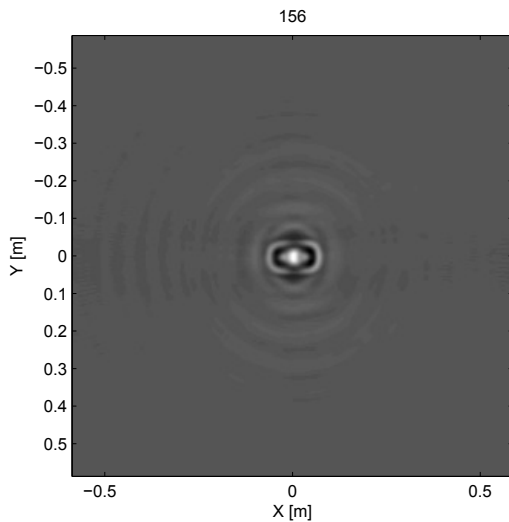


Fig. 6. Processed image of object illustrated in Figure 5 where the edges are clearly visible.

center of the rotating platform at coordinates (0,0). This was to see if an object within another object could be detected. Figure 8 shows the processed image of the object under test, and it is easy to spot the wooden box with the aluminum foil covered box inside. The center coordinates of the object and the size are correctly displayed. The squares inside of the box are due to constructive interference from waves reflected in corners [10].

The third object under test was a plastic egg tray from an egg boiler with three uncooked eggs in it shown in Fig. 9. In addition to placing the eggs 200 mm closer to the antenna, the eggs were placed 50 mm off the center line at $x=0$ to see if the image would capture the small displacement. Figure 10 shows the processed image of the object under test. The processed image clearly shows the three eggs located at (0.05,0.2), which corresponds with the initial state since the eggs were placed a bit off the center line.

Inspecting the results from these measurements, it is clear



Fig. 7. Test setup 2 - Wooden box on rotating platform with an aluminum foil covered cardboard box inside.

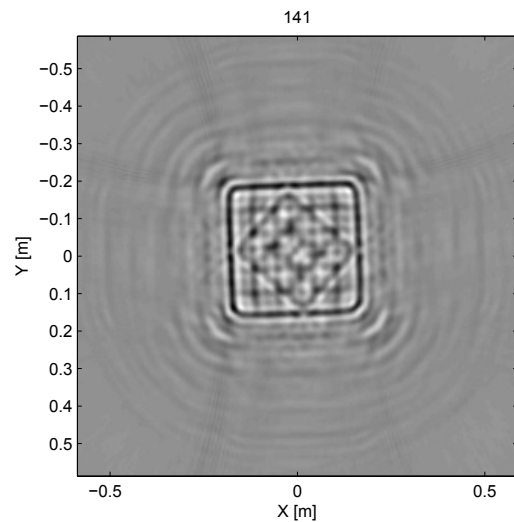


Fig. 8. Processed image of object illustrated in Figure 7 where both the outer and inner objects are clearly visible. The regular patterns are due to constructive interference.

that a two dimensional image of the cross section of an object under test can be acquired using the ISAR technique, even with quite complex objects. The results show that the initial position of the object is correctly displayed in the coordinate system, and the size of the objects matches the physical dimensions of the objects. The results also show that it is possible to both detect objects of different materials, not only metal or good conductors (reflectors), and to see through matter. The end results are calculated using measurement data that has room for improvement, which could enhance the image quality if addressed:

- A more accurate measurement of the components' coordinates in the set-up.
- Antenna position. What is the exact location of the antennas (x_t, y_t) and (x_r, y_r) ? The centers of the coax input/output from antennas were chosen as antenna position.



Fig. 9. Test setup 3 - Plastic egg dispenser with three raw eggs on rotating platform.

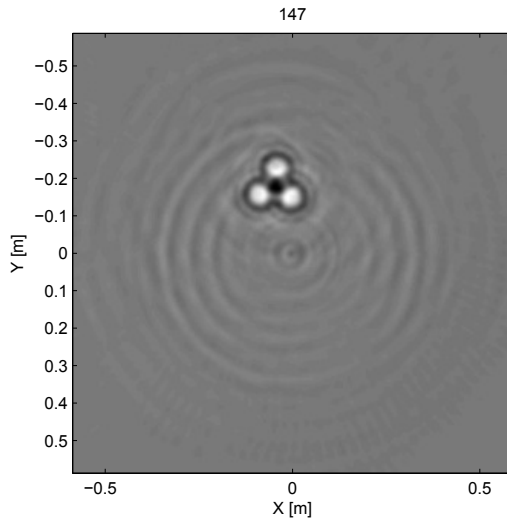


Fig. 10. Processed image of object illustrated in Figure 9 where the three eggs are clearly visible.

- Uneven steps of the step motor meaning that the angle used in DSP calculations may differ from the actual angle which is:

$$\frac{360^\circ}{288 \text{ steps}} = 1.25^\circ * \text{step}^{-1}$$

- Only single clutter map subtraction was used.
- More static environment (a person was present during measurements)
- Streaming frames from radar instead of one shot to keep same temperature and therefore delay for each frame [11].

The Nanoscale Impulse Radar has been characterized through tests and measurements and the main parameters of the system are listed in Tab. I.

V. CONCLUSION

In this paper a complete impulse based radar transceiver integrated on a single chip has been presented and utilized

Parameter	Value	Unit
TX Pulse Generator	1 st order Gaussian	NA
TX bandwidth unfiltered (-10dB cutoff)	0.85 - 9.6	GHz
TX nominal output power	-19	dBm
TX maximum PRF	100	MHz
RX sensitivity	-95	dBm
RX input threshold resolution	13	bits
RX equivalent sampling rate	39	GS/s
RX frame offset step	4	mm
Supply voltage core/IO	1.2/2.5	V
Power consumption (Max utilization)	113	mW

TABLE I
MEASURED KEY SYSTEM CHARACTERISTICS

for 2-D ISAR imaging. A resolution in the millimeter-range is achieved due to a combination of high bandwidth of the transmitted pulses and a very high speed Continuous Time Binary Valued (CTBV) strobed sampler. The ISAR technique was used to generate 2-D images of different objects using one single stationary radar system. The measured results support the related theory and confirm that the single chip Nanoscale Impulse Radar is a viable and low-cost alternative to more complex imaging in application areas such as airport luggage and full-body scanning, due to its penetration, positioning and precision properties.

ACKNOWLEDGMENT

The authors would like to thank the Research Council of Norway for funding this research through the grant agreement No. 193229.

REFERENCES

- [1] S. Gezici, "Theoretical limits for estimation of periodic movements in pulse-based uwb systems," IEEE J. Sel. Topics Signal Process., vol. 1, no. 3, pp. 405-417, Oct. 2007.
- [2] S. Venkatesh, C. Anderson, N. Rivera, and R. Buehrer, "Implementation and analysis of respiration-rate estimation using impulse-based uwb," in Proc. IEEE Military Communications Conf., 2005, vol. 5, pp. 3314-3320.
- [3] N. V. Rivera, S. Venkatesh, C. Anderson, and R. M. Buehrer, "Multitarget estimation of heart and respiration rates using ultra wideband sensors," presented at the XIV Eur. Signal Processing Conf., Florence, Italy, Sep. 4-8, 2006. [Online]. Available: <http://www.eurasip.org/Proceedings/Eusipco/Eusipco2006/papers/1568987536.pdf>.
- [4] T. E. McEwan, "Ultra-Wideband radar motion sensor," U.S. Patent 5361070, Nov. 1, 1994.
- [5] S. Azevedo and T. E. McEwan, "Micropower impulse radar," IEEE Potentials, no. 2, pp. 15-20, Apr./May 1997.
- [6] E. M. Staderini, "Uwb radars in medicine," IEEE Aerosp. Electron. Syst. Mag., vol. 17, no. 1, pp. 13-18, Jan. 2002.
- [7] H. A. Hjortland, D. T. Wisland, T. S. Lande, C. Limbodal, and K. Meisal, "Thresholded samplers for UWB impulse radar", Proceedings of Circuits and systems, 2007. ISCAS 2007. IEEE International Symposium on, 27-30 May 2007, New Orleans, LA, pp. 1210-1213.
- [8] http://en.wikipedia.org/wiki/Synthetic_aperture_radar, March, 2012.
- [9] M. Bury, Y. Yashchysyn and J. Modelski, "A Simple Approach for Elimination of Fixed Object's Reflections in UWB Imaging System," EUROCON 2007 The International Conference on "Computer as a Tool", Warsaw September 9-12.
- [10] http://www.met.rdg.ac.uk/clouds/maxwell/corner_reflector.html, March, 2012.
- [11] Novelda NVA-R640 Development Kit User Guide, March, 2012.

A.2 Localization and Communication

Paper 6 Power-efficient CTBV symbol detector for UWB applications

S. Sudalaiyandi, M. Z. Dooghabadi, T. A. Vu, H. A. Hjortland, Ø. Nass, T. S. Lande, and S. E. Hamran. “Power-efficient CTBV symbol detector for UWB applications”. In: *Ultra-Wideband (ICUWB), 2010 IEEE International Conference on*, pp. 1–4, Nanjing, Sep. 2010. doi:10.1109/ICUWB.2010.5614422. © 2010 IEEE. Reprinted with permission.

My Contributions to This Paper

Together with my supervisor I worked out the idea of CTBV symbol detectors for communication explored in this paper. A prerequisite for using delay lines with long time constants is calibration. I was active in working out the calibration solution required. I was also involved in writing the paper. My SPI solution as well as my microcontroller solution is used.

Power-Efficient CTBV Symbol Detector for UWB Applications

Shanthi Sudalaiyandi, Malihe Zarre Dooghabadi, Tuan-Anh Vu
Håkon A. Hjortland, Øivind Næss, Tor Sverre Lande and Svein Erik Hamran
Dept. of Informatics, University of Oslo, Norway
Email: shanthi@ifi.uio.no

Abstract—In this paper, we present a novel symbol detector architecture using time-domain running cross-correlation intended for impulse radio UWB communication. A bold new perspective on system design processing (CTBV) completely in continuous time is proposed and an implementation is carried out in 90 nm TSMC low power process technology. By exploring CTBV coding and unique combinational circuits for signal processing operations, a fast and power efficient solution is found to be feasible for clockless CMOS implementation.

I. INTRODUCTION

Almost a decade after the FCC release of the unlicensed ultra wideband (3.1 GHz–10.6 GHz), widespread use of this UWB band is still pending. Even with several commercial efforts [1][2][3] for single chip CMOS UWB solution, no real breakthrough has been made. Although several efforts for communication standards have been made [4], no efficient solution in standard CMOS is available. In spite of potential high performance with additional functionality like high-precision localization, power efficient, competitive CMOS solutions are hard to make. The combination of microwave frequencies with ultra wideband signal processing is a major challenge even with nanometer technology.

The major obstacle towards a power efficient UWB radio solutions are complicated channel estimation, accurate synchronization and high resolution clocks [5]. Due to the large bandwidth of UWB, the channel is extremely frequency sensitive and the received signal contains a large number of resolvable multipath components. Hence require RAKE or correlation receivers to take full advantage of the available signal energy. Because of complexity constraints, a RAKE receiver processes only a subset of the total number of received multipath components [6]. To mitigate the issues of the correlator receivers, different configurations have been proposed [6][7][8]. Most implementations sacrifice performance for low complexity operation and lacks a power efficient CMOS implementation aiming towards a single chip solution.

At the Nanoelectronics group, Dept. of Informatics, we have brought forward a different way of designing impulse radio (IR-UWB) based systems using pure time-domain signal processing. The Continuous-Time Binary-Value (CTBV) approach is simply based on inherent, process-dependent gate delays. By avoiding clocks, power efficient solutions are feasible in standard CMOS. Although process variations pose significant challenges, high-speed and power-efficient CMOS

implementations have been demonstrated [9][10]. A similar approach is proposed in [11] called as Continuous-Time Discrete Amplitude (CTDA). The continuous time sampling systems compared to the discrete time avoids quantization noise and clocks [12]. Also continuous time systems has fewer distortion components at frequencies within the band, input activity dependent power dissipation and fast response to sudden input variations [12].

In our current work, we explore CTBV coding for power efficient impulse radio (IR-UWB) receivers. Impulse radio based communication systems are characterized by simple and power efficient transmitters, while receivers are complicated and hard to make. These receivers are often called *correlating receivers* because template matching is required [13]. Running cross-correlation is often power demanding. However, simple and power-efficient solutions are feasible using CTBV coding. To compensate for the process mismatch a simple initial calibration is required.

A novel symbol detector in time-domain implementation is proposed in this paper, which overcomes the shortcomings of the conventional correlator receivers without any increase in the complexity of the receiver structure using CTBV approach and signal processing operations using unique combinational circuits. The symbol detector is based on cross-correlation symbol-detection, storing the incoming signal using the delay elements (equivalent to the chip rate) and performing cross-correlation with the reference template. The advantages of the proposed symbol detector is the following: (1) *Power efficient CMOS implementation*: We can implement efficient cross-correlator using simple combinatorial logic instead of the complex multi-gigahertz DSPs, giving low cost and power efficient solution. (2) *CTBV domain*: The proposed symbol detector uses CTBV signal coding completely without power-demanding clocks. CTBV coding combined with suitable binary signal processing solutions give the speed of analog signal processing and some of the advanced processing capabilities available with digital signal processing [14]. The computational efficiency is shown with the high speed sampler of an impulse radar in CTBV implementation with sampling rate exceeding 30 GHz[9]. (3) *Delay calibration*: Suitable timing is accomplished by cascading several unit-delay elements (inverters) combined with calibration. The performance of the CTBV solutions is only limited by the gate delays and scales well with process technology.

II. UWB SYMBOL DETECTOR ARCHITECTURE

The UWB symbol detector architecture as shown in Fig. 1 consists of a pulse extender, tunable delay elements, template register, correlator, continuous time counter, thresholder and the data readout.

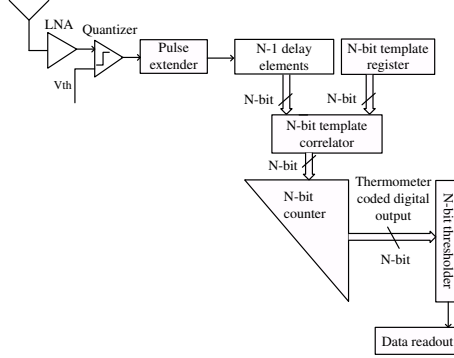


Fig. 1. Block diagram of the receiver.

The proposed symbol detector is based on Code Division Multiple Access (CDMA) coding, which uses unique spreading codes. The principle of the symbol detector is as illustrated in Fig. 2. The incoming pulses are amplified through a LNA and passed to a single-bit quantizer. According to the defined threshold voltage, the 1-bit quantizer converts the incoming pulse into CTBV signals. For increased detectability the quantizer output is pulse extended using an adapted digital circuit. The incoming bit-sequence is stored in a delayline and correlates with the template in continuous time. In CTBV coding, correlation refers to the matching of incoming '1's with the input template. The correlated output are counted and coded as a thermometer coded digital output. The matching result may then be thresholded with an appropriate detection level using the thresholder. The different processing blocks are explained, as follows.

A. Pulse Extender

The output of the quantizer may be very short pulses (<100 ps duration). For proper signal processing time, a simple pulse extender is inserted in the signal path. The schematic of the pulse extender as shown in Fig. 3b, expand the pulse width of the incoming symbol by prolonging the rise time of the inverter using a weak PMOS. The RC time constant is tuned according to the desired pulse width and is recovered by a schmitt trigger. The slow rise time is prone to jitter and process variations, but can be compensated by tuning, to some extent. The pulse extender can be stretched to a maximum value of ≈ 3 ns.

B. Delay Line

The delay line consists of $N-1$ delay elements of delay time T as shown in Fig. 3a. The cascaded delay elements provides a history of the incoming signal. By tapping the delay line in suitable intervals matching the chip-rate of the transmitted symbol, the symbol code should appear simultaneously as a

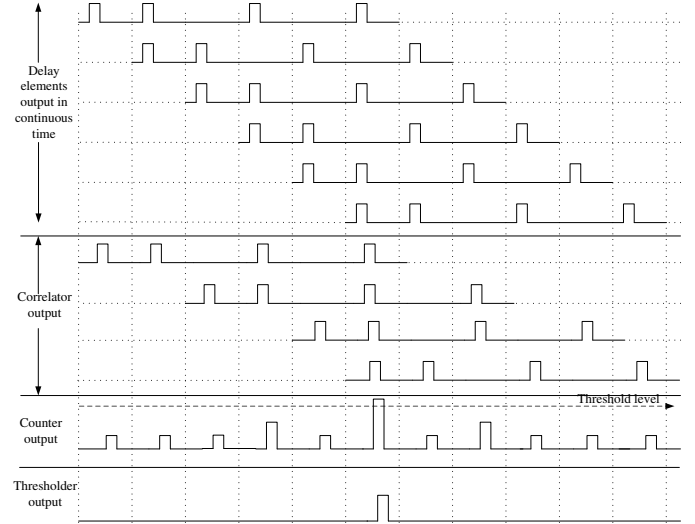


Fig. 2. The symbol detection (Example-110101).

bit-pattern. As the decision time for the pulse is limited by the pulse extender, the mismatch between the taps must be tuned. We have developed a tunable delay element matching the chip-length. The tuning or calibration procedure is achieved by programming each delay element through coarse-tune and fine-tune register settings using a SPI controller. The coarse delay can be tuned to a maximum value of 16 ns with a resolution of 0.5 ns and the fine delay can be tuned to a maximum value of 1.6 ns with a fine resolution of 0.05 ns. The delay time T of a single element is ≈ 21.6 ns (fixed (4 ns) + fine tune (1.6 ns) + coarse tune (16 ns)). The coarse delay element uses the delay structure as shown in Fig. 3c, which consists of cascaded NMOS and PMOS structures, giving increased gate delay with minimal area overhead. This delay structure is one fifth of a coarse delay element which is ≈ 0.1 ns. An average inverter gate delay is ≈ 15 ps, which is limited by the process technology. For unit delays with chip length of ≈ 20 ns we expect 0.05 ns resolution to be sufficient.

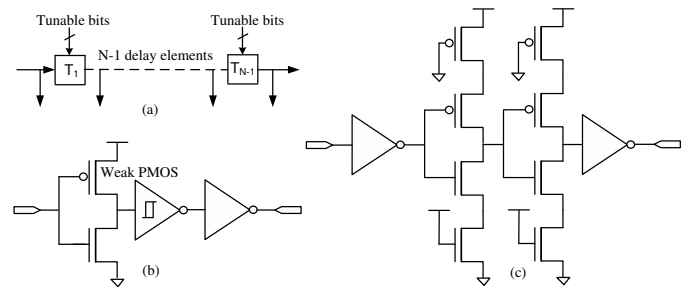


Fig. 3. (a) Overview of the delay line. (b) Schematic of a pulse extender. (c) Schematic of a 0.1 ns delay structure.

C. Correlator

The taps from the delay line are correlated with the input template using the N -bit correlator in continuous time. Since the symbol detector is designed for detecting occurrences of

pulses (bit '1') in the received symbol, the correlated output consists of a sequence of bits where matches are indicated by '1'. The remaining computation is to compute number of '1's at the correlator output.

D. Continuous Time Counter

The novel continuous time N-bit counter is implemented using 2-input multiplexers, arranged in columns as shown in Fig. 4. As more matches occur, '1'-bits are routed to the outputs. The number of matches between the incoming bit-sequence and the template register will continuously appear at the counter output as a thermometer-coded value. The inputs into the counter has been skewed in time to compensate for the difference in propagation between the different counter inputs. The pulse stretching just gives us some safety margins.

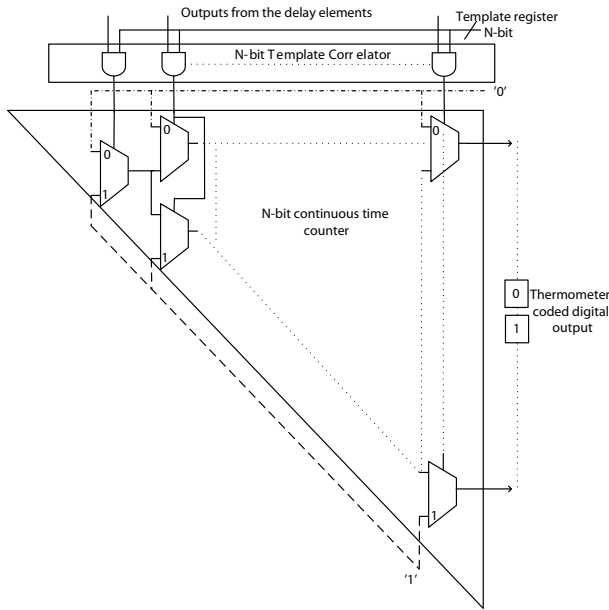


Fig. 4. Block diagram of an N-bit continuous time counter.

E. N-bit Thresholder and Data Readout

The output from the continuous time counter is analyzed through the block shown in Fig. 5. The matching thresholder block consists of the N-bit threshold register, N-bit thresholder, glitch remover circuit and the output readout.

1) *N-bit Thresholder*: The N-bit threshold register can be programmed through the SPI controller according to the required matching for symbol detection as shown in Fig. 5. Depending on symbol code, the appropriate threshold may be set. The N-bit thresholder logic compares the thermometer-coded output from the cross-correlation with the threshold register setting for continuous symbol detection. The symbol may be sent multiple times for improved symbol recovery (adds the processing gain). For a given matching threshold set by the threshold register, we may continuously check for match by simply AND-ing with the counter output and then OR-ing to get the symbol detection.

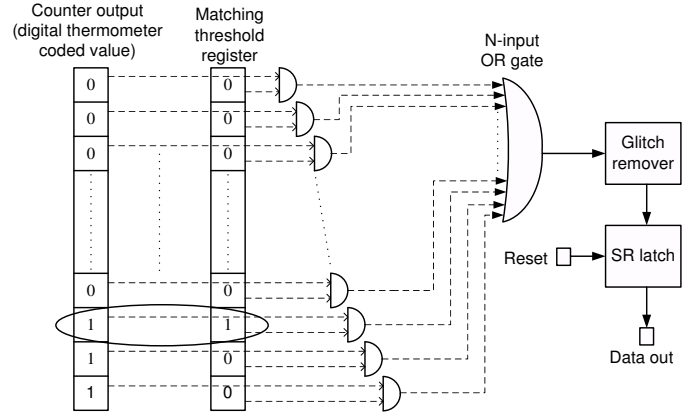


Fig. 5. Block diagram of the thresholder.

2) *Glitch Remover*: To prevent any undesired glitches at the output of thresholder circuit, the correlated output is passed to a glitch remover circuit as shown in Fig. 6. The circuit is based on the pulse extender circuit and is simply a combination of two pulse extenders.

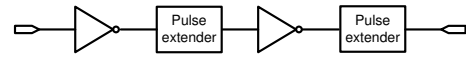


Fig. 6. Glitch remover.

3) *Output Readout*: The occurrence of symbol detection in continuous time is analyzed using the SR latch.

III. SIMULATION RESULTS

A 32-bit symbol detector was designed and implemented using the TSMC 90 nm low power process technology. The top level layout of the symbol detector is as shown in Fig. 7. All the complex signal processing operations are performed using basic combinational logic. In order to achieve high data rates, we can increase the chip-rate but in practical systems we are limited by the inter-symbol interference. We designed for a chip duration of 10 ns with pulse extended to 2 ns. A symbol size of 32 bits were used since the choice of delay taps play a significant role in the system performance [6]. Post-layout simulations on different blocks using Cadence is performed for validation of the proposed design.

The short signal from the quantizer output are stretched by the pulse extender circuit as shown in Fig. 8a. The received bit-sequence is stored in the 32-bit delay line. The transient simulation results of a single delay element with all the enable bits tied high (maximum length) is shown in Fig. 8b. Each delay element is designed such that they are tunable up to 35 ns. This is important to have enough decision time to sample the incoming pulse in the continuous time. For improved readability, the transient simulations of the counter for a 5-bit symbol representing the thermometer coded digital value is as shown in Fig. 8d for the corresponding input sequence as shown in Fig. 8c. The maximum delay of this counter are approx. 2.5 ns. The occurrence of the bit '1' in the thermometer coded counter outputs are detected with the

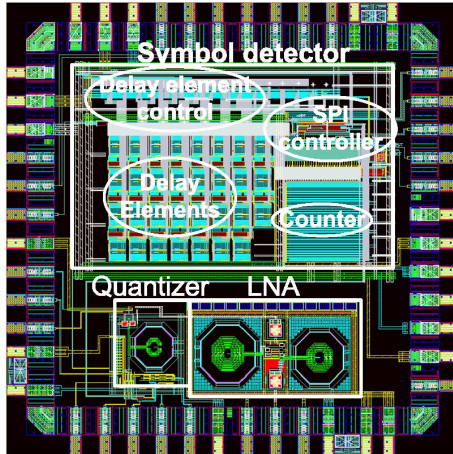


Fig. 7. Layout of the receiver.

appropriate setting in the threshold register as shown in Fig. 5. The detected signal is validated through the glitch remover circuit to get rid off any pulses shorter than 100 ps as shown in Fig. 8e. The glitch free detected signal will set the SR latch and read out for further analysis.

IV. CONCLUSION

In this paper we are proposing a novel continuous-time symbol detector for impulse radio receivers suitable for power efficient implementation in standard CMOS. A running cross-correlation is used for symbol recovery. Post-layout simulation validates the proposed architecture. The performance will be verified by measurements of a test chip fabricated in 90 nm CMOS using TSMC technology.

ACKNOWLEDGMENT

The research is being sponsored by Norwegian Research Council project number 187857/S10 and is carried out at the Nanoelectronics group, Department of Informatics, University of Oslo, Norway.

REFERENCES

- [1] [Online]. Available: <http://www.staccatocommunications.com/>
- [2] [Online]. Available: <http://www.pulselink.net/>
- [3] [Online]. Available: <http://www.timedomain.com/>
- [4] [Online]. Available: <http://iee802.org/15/index.html>
- [5] Y. Wang, G. Leus, and A.-J. van der Veen, "Digital receiver design for transmitted reference ultra-wideband systems," *EURASIP Journal on Wireless Communications and Networking*, vol. 2009, no. 13, 2009.
- [6] J. D. Choi and W. E. Stark, "Performance of ultra-wideband communications with suboptimal receivers in multipath channels," *IEEE Journal on Selected Areas in Communications*, vol. 20, no. 9, pp. 1754–1766, Dec. 2002.
- [7] Y.-L. Chao and R. A. Scholtz, "Ultra-wideband transmitted reference systems," *IEEE Transactions on Vehicular Technology*, vol. 54, no. 5, pp. 1556–1569, Sep. 2005.
- [8] A. T. Phan, R. Farrell, M.-S. Kang, and S.-G. Lee, "Low-power sliding correlation cmos uwb pulsed radar receiver for motion detection," in *IEEE International Symposium on Circuits and Systems*, Jun. 2009, pp. 1541–1544.
- [9] H. A. Hjortland, D. T. Wisland, T. S. Lande, C. Limbodal, and K. Meisal, "Thresholded samplers for uwb impulse radar," in *IEEE International Symposium on Circuits and Systems, 2007. ISCAS 2007.*, 2007.
- [10] Øyvind Dahl, H. A. Hjortland, T. S. Lande, and D. T. Wisland, "Close range impulse radio beamformer," in *2009 IEEE International Conference on Ultra-Wideband*, Sep. 2009, pp. 200–204.
- [11] Y. Li, K. Shepard, and Y. Tsividis, "A continuous-time programmable digital fir filters," *IEEE Journal of Solid-State Circuits*, vol. 41, no. 11, pp. 2512–2530, Nov. 2006.
- [12] B. Schnell and Y. Tsividis, "A continuous-time adc/dsp/dac system with no clock and with activity-dependent power dissipation," *IEEE Journal of Solid-State Circuits*, vol. 43, no. 11, pp. 2472–2481, Nov. 2008.
- [13] J. D. Taylor, *Ultra-Wideband Radar Technology*. CRC Press, 1995.
- [14] H. A. Hjortland and T. S. Lande, "Ctbv integrated impulse radio design for biomedical applications," *IEEE Transactions on Biomedical Circuits and Systems*, vol. 3, no. 2, pp. 79–88, Apr. 2009.

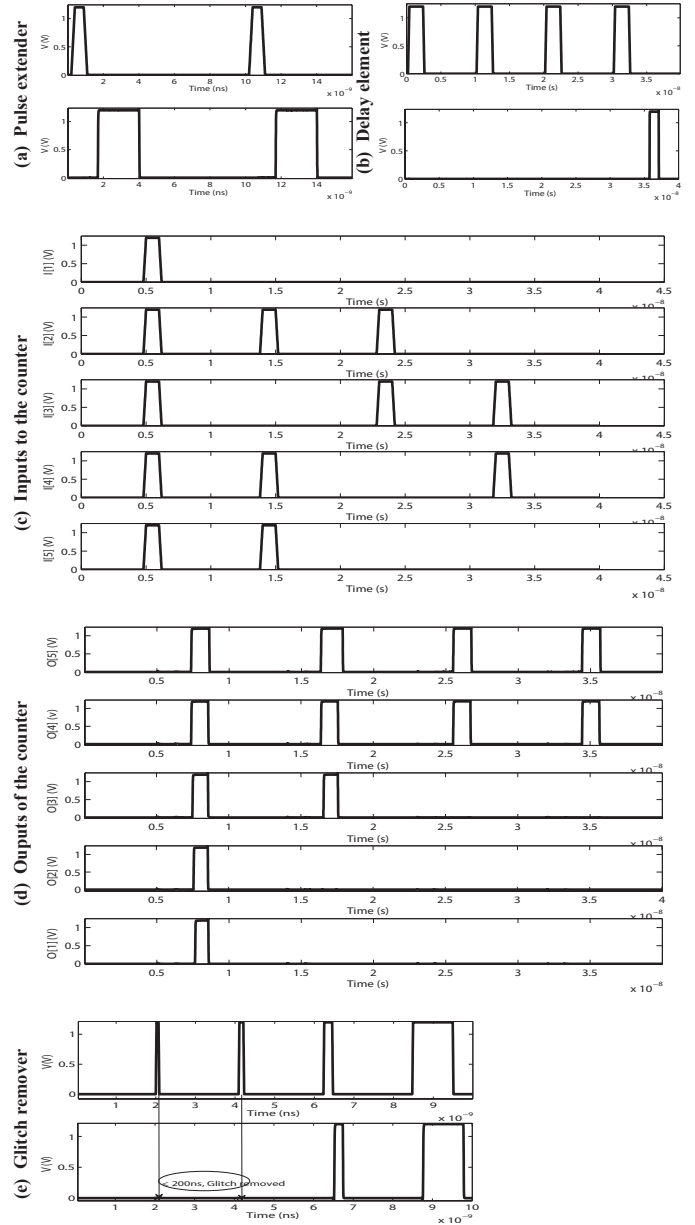


Fig. 8. Simulation results of (a) pulse extender. (b) Delay element. (c) Counter inputs. (d) Counter outputs. (e) Glitch remover.

Paper 7 Continuous-time single-symbol IR-UWB symbol detection

S. Sudalaiyandi, T.-A. Vu, H. A. Hjortland, O. Nass, and T.-S. Lande. “Continuous-time single-symbol IR-UWB symbol detection”. In: *SOC Conference (SOCC), 2012 IEEE International*, pp. 198–201, Niagara Falls, NY, Sep. 2012. doi:10.1109/SOCC.2012.6398347. © 2012 IEEE. Reprinted with permission.

My Contributions to This Paper

This paper presents a continuous-time RAKE receiver or symbol detector. I came up with the design for much of the system and many of the circuits. One of the important points in the paper is to improve the control of the pulse widths through the system by designing the delaylines more robust against process variations and by using active pulse width controller blocks. I helped design both of these solutions and the algorithm used for tuning delays and pulse widths. I also made the symbol code used here, I designed the continuous-time counter together with my supervisor, and made the SPI and microcontroller system which is used to control the chip. Furthermore, I created the symbol detection principle plot, and I also helped review the paper.

Continuous-Time Single-Symbol IR-UWB Symbol Detection

Shanthi Sudalaiyandi, Tuan-Anh Vu, Håkon A. Hjortland,
Øivind Næss, and Tor Sverre Lande

Dept. of Informatics, University of Oslo, Norway, Email: shanthi@ifi.uio.no

Abstract—RAKE based receivers are the most challenging receiver due to its hardware implementation complexity. To build a cross-correlating RAKE receiver with on-off keying pose significant design challenges due to multipath components. Also process variations in 90 nm CMOS technology impose significant challenges in the clockless system. In this paper, a simple coherent RAKE receiver using single-bit correlation in continuous time is reported. A working chip with fully integrated digital symbol detection without any clocks, ADC and filters for simple and power efficient CMOS implementation. Optimal design of the delay elements in the tapped delayline and the use of pulse width controller blocks (PWC) along the delayline allows narrow pulses through the system. The optimized delayline combined with the continuous-time high speed counter and the match thresholder constitute a complete continuous-time RAKE receiver in CMOS. Initial chip measurement shows that a data rate of 1.5 Mbit/s with a power consumption of 0.78 mW was possible.

I. INTRODUCTION

Even though viable UWB systems are not in widespread use, there are many published papers which show great potential towards short range communications requiring low power and low data-rate [1]. Ultra-wideband impulse radio (IR-UWB) is a promising technology due to its good resolution. According to [2], channel estimation, accurate synchronization and high resolution clocks forms the bottleneck to realize a power efficient UWB radio solutions. UWB signals require RAKE or correlation receivers to take full advantage of the available signal energy over a large bandwidth. Because of complexity constraints, a RAKE receiver processes only a subset of the total number of received multipath components. To mitigate the issues of correlator receivers, different configurations have been proposed [3]. Most implementations sacrifice performance for low complexity operation and lacks a power efficient CMOS implementation aiming towards a cost efficient single chip solution. Even if the receiver complexity is overcome by digital techniques, current solutions require an ADC with high sampling frequency to sample and quantize the received signal [3].

Continuous-Time Binary-Value (CTBV) signal processing for IR-UWB systems has been developed [4] which is simply based on inherent CMOS process-dependent gate delays combined with a coarse quantizer or a single-bit ADC. By avoiding clocks, power efficient solutions are feasible in standard CMOS. The performance of the CTBV solution is limited by the gate delay and improves with technology scaling. Although process variations pose significant challenges, high-speed and power-efficient CMOS implementations have

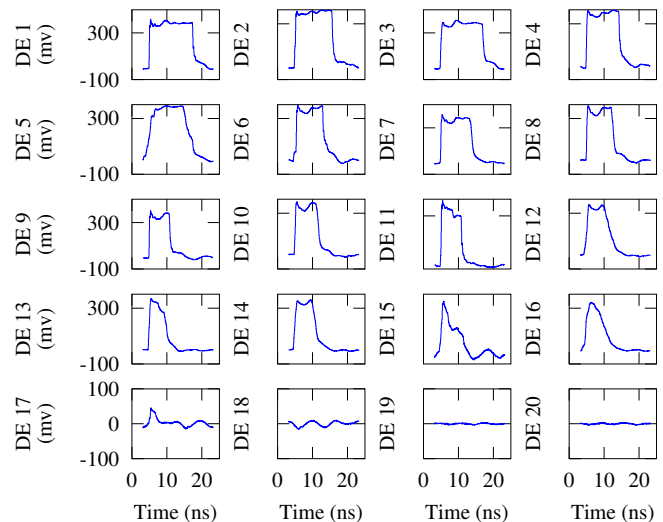


Fig. 1. Pulse disappearing across the delayline.

been demonstrated using unique combinational circuits [5]. A similar approach is proposed in [6] named Continuous-Time Discrete Amplitude (CTDA). The continuous time sampling systems compared to the discrete time avoids quantization noise and clocks [7]. Also continuous time systems has fewer distortion components at frequencies within the band, input activity dependent power dissipation and fast response to sudden input variations [7].

A symbol detector using CTBV coding in time-domain implementation may provide power-efficient solution in conventional correlator receivers with minor increase in receiver complexity. The symbol detector in [8] is based on CTBV coding symbol-detection, storing the incoming signal in a suitable delayline (matching to the chip interval) continuously performing cross-correlation with the reference template. The major challenge of continuous-time symbol detector is the process variation and mismatch, due to which short pulses disappear in the long delayline. In [8], measurement results showed around 1 ns of pulse shortening per delay element (DE) throughout the delayline. Given the fact that the IR-UWB pulses can have components less than 1 ns, the incoming pulses in the symbol may disappear in the delayline as shown in Fig. 1. Disappearance of pulses in the delayline posed a significant challenges to realize a CTBV symbol detector on chip. Also since CTBV system does not use any clocks for

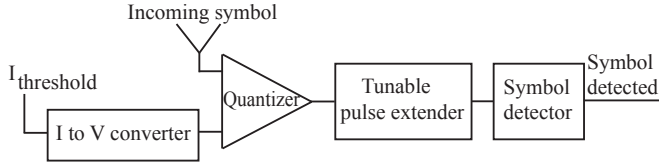


Fig. 2. Block diagram of the continuous-time rake receiver.

symbol detection, process variation and mismatch are very problematic leading to synchronization issues.

In this paper, we explore CTBV coding for power efficient impulse radio receivers and the focus is on continuous-time symbol detector. We have shown that with careful design of the delay elements (DE) combined with proper initial calibration have reduced the process variation impact on the delayline. Also we have designed pulse width controllers (PWC) in the delayline maintaining pulse width throughout the delayline. The rest of the paper is organized as follows. II gives an overview of the continuous-time rake receiver. III gives an detailed description of the symbol detector architecture and the design methodology to mitigate the issue of pulse shrinking. Measured results of the chip are explained in IV.

II. CONTINUOUS-TIME RAKE RECEIVER

The block diagram of a continuous-time CTBV receiver is as shown in Fig. 2. A single ended 1-bit quantizer with tunable threshold setting is at the receiver front-end. The quantizer is essentially a thresholder continuously comparing the incoming signal to a threshold voltage, giving a digital output which is a CTBV coded signal [9]. The optimal threshold value for the quantizer is a trade-off between noise and the shorter pulse width signals, which is an iterative process to fix. Also the dependency of temperature and process variation complicates this process. Hence the threshold is set high enough to get rid most of the noise and other ringing effects, and with the help of pulse extender the signal is reconstructed with sufficient pulse width to pass through the deep digital logic of the symbol detector. The pulse extended CTBV signal is then cross-correlated with a symbol template to detect the incoming symbol in the symbol detector block. Thus performing both threshold detection and the correlation detection on the incoming signal, we are able to detect the symbol.

III. CONTINUOUS-TIME SYMBOL DETECTOR ARCHITECTURE

The continuous-time symbol detector architecture (Fig. 3) is designed for a symbol length of 32, which consists of delayline, correlator, counter and match thresholder. The quantized output may be very short pulses and hence pulse extended using an adapted digital circuit. The basic working principle of the continuous-time symbol detector is as shown in Fig. 4. The CTBV pulse sequence is stored in a delayline and correlated with the template in continuous-time. The symbol detector is based on on-off keying (OOK) and hence correlation refers to the matching of incoming '1's with the input template. The

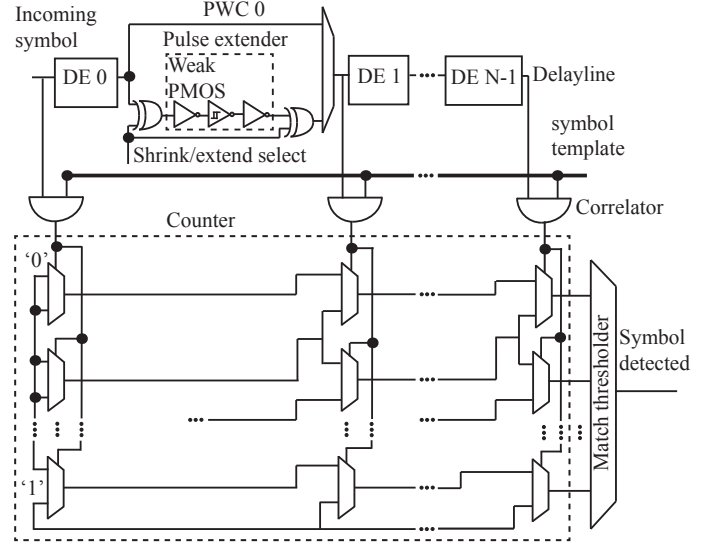


Fig. 3. Schematic of the symbol detector.

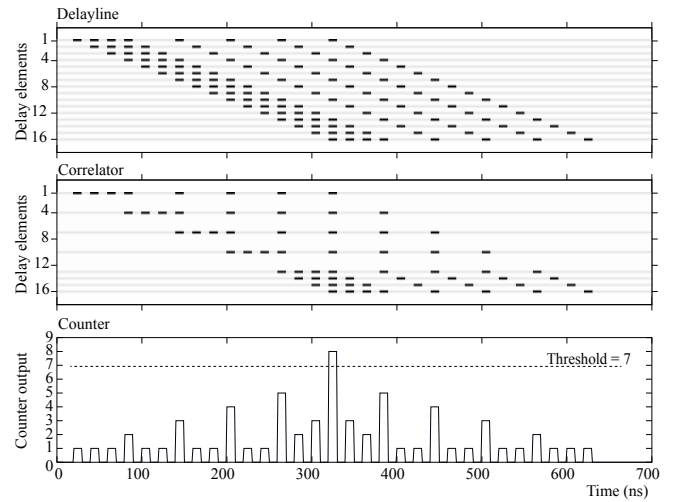


Fig. 4. Principle of symbol detection for symbol pattern 1111001001001001.

correlated output are counted and coded as a thermometer coded digital value using the high speed counter. The counter output may then be thresholded with an appropriate detection level using the match thresholder. By setting a threshold value in the match thresholder, we can efficiently detect the number of pulses in the incoming symbol.

A. Delayline

By tapping the delayline in suitable intervals matching the chip-interval of the transmitted symbol, the symbol will appear simultaneously as a pulse pattern. The delayline consists of 31 delay elements (DE) to process the incoming 32-chip symbol. Each delay element consists of coarse and fine tune, wherein coarse delay element is designed using cascaded NMOS and PMOS structures and fine tune delay elements are designed using the minimum sized inverters (limited by the process technology). The process dependent variation, mismatch, and

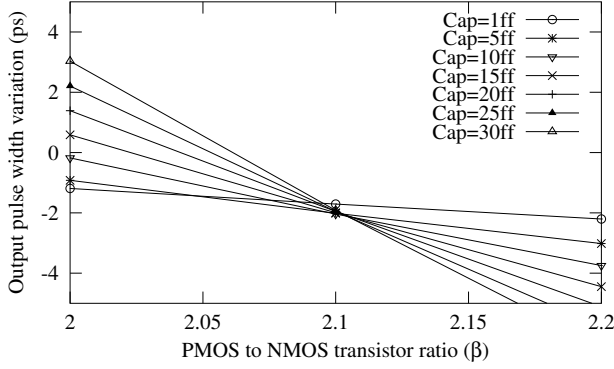


Fig. 5. Optimal design of delay element.

jitter impacts the delayline and thus the most critical block in the RAKE based systems. The above mentioned impact in the delay element causes variation both in the pulse position and also the pulse width of the incoming symbol.

Any mismatch in the delay element is calibrated to match the chip-interval of the incoming symbol. The tuning or calibration procedure is achieved by programming each delay element through the coarse-tune and fine-tune register settings using a SPI microcontroller. With respect to the first delay element, the second delay element is tuned such that they have maximum overlap in pulse width between them. With the first and second correlators enabled, we only have 2 symbols aligned in time. By setting the match threshold to 2, pulse is detected at the output if the first and the second delay elements are tuned properly in time. In the similar way, other delay elements are also tuned with respect to the first delay element. In-depth details of the calibration is beyond the scope of this paper and will be published in future.

The delayline consists of long chain of multiplexers and inverters, which forms asymmetrical capacitive loading to the signal path. The inverters in the delay element are hence carefully designed such that their pulse width variation are the same irrespective of the different capacitive loading as shown in Fig. 5. Thus any shrinking in the pulse width is balanced out by the inverting and non-inverting logic present in the delayline. Further variation in the pulse width of the incoming symbol can be tuned by the pulse width controller blocks (PWC) present in the delay element. As variation and mismatch causes an uncorrelated effect on the delayline, the pulse width of the incoming pulse after each delay element is either extended or shrunk using the PWC.

B. High Speed Correlation

There are different analog and digital correlators available that are suitable for a wideband correlation. Straight forward implementations of the digital correlators are very power hungry demanding fast DSPs to perform several giga multiplications and adds per second [10]. On the other hand, analog implementations are very sensitive to noise and process variations [11]. In this paper, the symbol detector uses a digital correlator based on a simple AND gate, which continuously

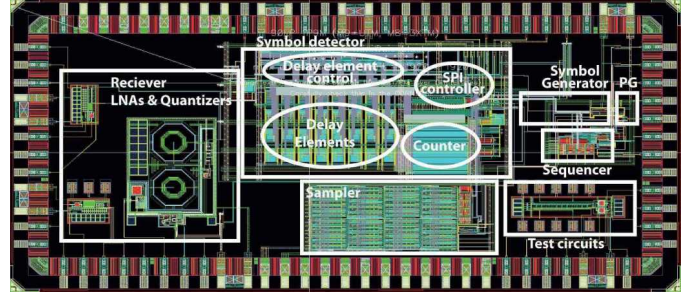


Fig. 6. Layout of the CTBV system.

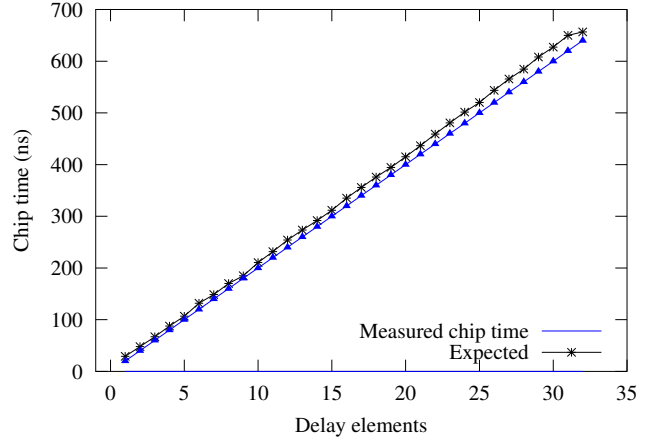


Fig. 7. Measured chip interval across each delay element.

correlates the delayline output and detects the number of pulses (number of correlation matches) in the received symbol. By avoiding fast clocking registers and parallelism, the correlators are very power-efficient and at the same time performs high speed correlation without any overhead. The detected number of pulses at the correlator output are continuously routed to the counter output as a thermometer-coded digital value, which simplifies the readout circuitry.

The match threshold consists of a multiplexer, which is programmed through the SPI controller. Depending on the detected number of pulses in the incoming symbol, the appropriate threshold may be set. The output circuitry consists of a continuous-time threshold sampler [5] and an integrated symbol generator which uses the Gaussian pulse generator [12] to generate the required symbol for transmission.

IV. MEASURED RESULTS

A continuous-time symbol detector was designed and fabricated using the TSMC 90 nm low power process technology. The top level layout of the CTBV system showing the symbol detector is as shown in Fig. 6. A symbol size of 32 chips were used since the choice of delay taps play a significant role in the system performance. Increasing the number of chips is a trade-off between the hardware implementation and the processing gain. In order to achieve high data rates, we can increase the chip-time but inter-symbol interference has a limit on this. Hence a chip-time of 20 ns with pulse width around 5 ns was

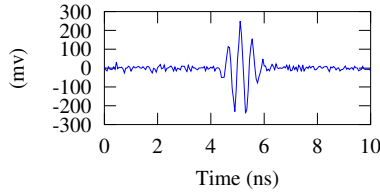


Fig. 8. Gaussian pulse.

used. With optimal design of delay elements along with the use of PWC blocks, we are able to see the pulses across each delay element without getting shrunk. This is obvious in Fig 7, which also indicates that most of the static variations in the chip time are fixed to some extent in the delayline through the calibration technique.

For the purpose of measurement, a symbol pattern based on gold code (auto-correlation = 16 and cross-correlation = 8) was transmitted as shown in Fig. 9(a). Each pulse in the symbol is a Gaussian pulse as shown in Fig. 8. More detailed description of this transmitter is beyond the scope of this paper and will be published in future. The correlator detects the presence of 16 pulses in the symbol and the counter counts this 16 pulses as a thermometer coded value. If the incoming symbol matches the symbol template and by setting the match threshold to 16 (maximum pulses), the symbol is detected and a pulse is seen at the output of the symbol detector as shown in Fig. 9(b).

The data rate $(1/((T_{chip} * N_s + T_{quiet}) * n_{repetition}))$ is calculated as 1.5 Mbit/s. T_{chip} is the chip interval and T_{quiet} (≈ 30 ns) is the assumed quiet period after every symbol transmission. $n_{repetition}$ gives the number of times symbol is transmitted. N_s corresponds to number of chips. The power consumption of the symbol detector is 0.78 mW and the energy consumed is around 15.6 pJ/pulse. Initial measured results shows that the processing gain increases the chance of symbol detection although it reduces the data rate. The coherency of the symbol detection is obtained since cross-correlation is running in continuous-time and the detection pulse will be issued immediately when matching occurs. This feature may be further explored for precision ranging in localization architectures.

V. CONCLUSION

A coherent CTBV symbol detection using running cross-correlation without any clocks at a data rate of 1.5 Mbit/s was designed in 90 nm CMOS technology and initial measured results are provided. With the careful design of delay elements and pulse width controller blocks, the pulse shrinking effect throughout the delay line has been reduced. The novel receiver architecture is a promising solution for IR-UWB receivers solving some of the fundamental implementation limitations of conventional RAKE architectures.

ACKNOWLEDGMENT

The research is being sponsored by Norwegian Research Council project number 187857/S10 and is carried out at the

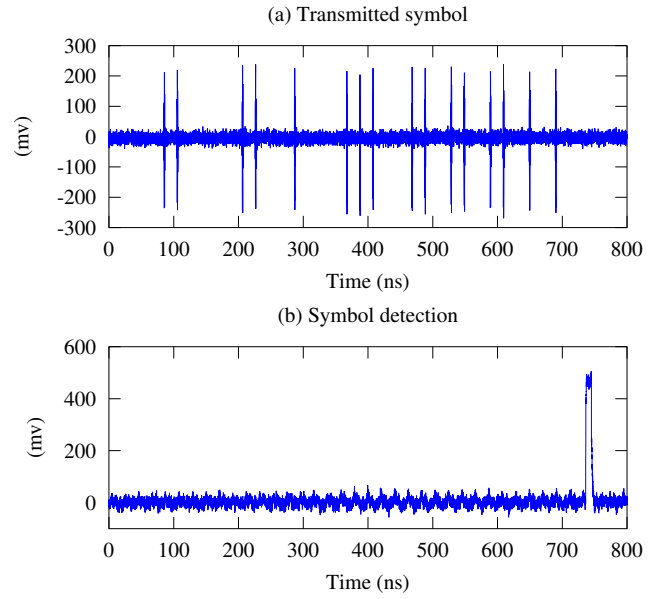


Fig. 9. (a) Transmitted symbol (1100001100100011100110110110101). (b) Continuous-time symbol detection.

Nanoelectronics group, Department of Informatics, University of Oslo, Norway.

REFERENCES

- [1] Y. J. Zheng, S. Diao, C. Ang, Y. Gao, F. Choong, Z. Chen, X. Liu, Y. Wang, X. Yuan, and C. H. Heng, "A 0.92/5.3nj/b uwb impulse radio soc for communication and localization," in *IEEE International Solid-State Circuits Conference*, feb. 2010, pp. 230–231.
- [2] W. Yiyin, L. Geert, and A. van der Veen, "Digital receiver design for transmitted reference ultra-wideband systems," *EURASIP Journal on Wireless Communications and Networking*, 2009.
- [3] Y. Chao and R. A. Scholtz, "Ultra-wideband transmitted reference systems," *IEEE Transactions on Vehicular Technology*, 2005.
- [4] H. A. Hjortland and T. S. Lande, "Ctbn integrated impulse radio design for biomedical applications," *IEEE Transactions on Biomedical Circuits and Systems*, pp. 79–88, apr. 2009.
- [5] H. A. Hjortland, D. T. Wisland, T. S. Lande, C. Limbodal, and K. Meisal, "Thresholded samplers for uwb impulse radar," in *IEEE International Symposium on Circuits and Systems*, may. 2007, pp. 1210–1213.
- [6] Y. W. Li, K. L. Shepard, and Y. P. Tsividis, "A continuous-time programmable digital fir filter," in *IEEE Custom Integrated Circuits Conference*, sep. 2005, pp. 695–698.
- [7] B. Schell and Y. Tsividis, "A continuous-time adc/dsp/dac system with no clock and with activity-dependent power dissipation," *IEEE Journal of Solid-State Circuits*, pp. 2472–2481, nov. 2008.
- [8] S. Sudalaiyandi, M. Z. Dooghabadi, T.-A. Vu, H. A. Hjortland, O. Naess, T. S. Lande, and S. E. Hamran, "Power-efficient ctbn symbol detector for uwb applications," in *IEEE International Conference on Ultra-Wideband*, sep. 2010, pp. 1–4.
- [9] T. A. Vu, S. Sudalaiyandi, M. Z. Dooghabadi, H. A. Hjortland, O. Naess, T. S. Lande, and S. E. Hamran, "Continuous-time cmos quantizer for ultra-wideband applications," in *IEEE International Symposium on Circuits and Systems*, jun. 2010, pp. 3757–3760.
- [10] M. Verhelst, W. Vereecken, M. Steyaert, and W. Dehaene, "Architectures for low power ultra-wideband radio receivers in the 3.1-5ghz band for data rates <10mbps," in *International Symposium on Low Power Electronics and Design*, aug. 2004, pp. 280–285.
- [11] S. Sriram, K. Brown, and A. Dabak, "Low-power correlator architectures for wideband cdma code acquisition," in *Conference Record of the 33 Asilomar Conference on Signals, Systems, and Computers*, 1999.
- [12] K. K. Lee, M. Z. Dooghabadi, H. A. Hjortland, O. Naess, and T. S. Lande, "A 5.2 pJ/pulse impulse radio pulse generator in 90 nm cmos," in *IEEE International Symposium on Circuits and Systems*, may. 2011.

Paper 8 Continuous-time high-precision IR-UWB ranging-system in 90 nm CMOS

S. Sudalaiyandi, H. A. Hjortland, T. Vu, O. Næss, and T. S. Lande. “Continuous-time high-precision IR-UWB ranging-system in 90 nm CMOS”. In: *Solid State Circuits Conference (ASSCC), 2012 IEEE Asian*, pp. 349–352, Kobe, Nov. 2012. doi:10.1109/ASSCC.2012.6570786. © 2012 IEEE. Reprinted with permission.

My Contributions to This Paper

This paper presents a high precision IR-UWB time of flight ranging system. I helped design much of the system and many of the circuits, such as the continuous-time symbol detector, the symbol transmitter, and calibration solutions. I helped getting the system up and running and ready for measurements, and I helped analyze and plot the data. I also helped review the paper.

Continuous-Time High-Precision IR-UWB Ranging-System in 90 nm CMOS

Shanthi Sudalaiyandi, Håkon A. Hjortland, Tuan-Anh Vu,
Øivind Næss, and Tor Sverre Lande

Dept. of Informatics, University of Oslo, Norway, Email: shanthi@ifi.uio.no

Abstract—Precision ranging is a key factor for localization in wireless sensor networks. Most reported ranging solutions are limited by clock frequency or clock synchronization. With the combination of continuous-time binary value (CTBV) technique and the impulse-radio ultra-wideband (IR-UWB) technology, we have a promising approach towards high precision positioning combined with communication. In this paper, we have presented a UWB communication solution with embedded ranging by measuring the symbol round trip time without a common reference clock. A first working clockless UWB ranging transceiver realized in 90 nm CMOS technology with fully integrated digital symbol detection including a running cross-correlator is implemented. Preliminary measured results shows that two transceivers in a master-slave configuration can estimate distance with centimeter precision (≈ 1.4 cm).

I. INTRODUCTION

Recently, ultra-wideband (UWB) technology has been studied for ranging and localization in wireless sensor networks (WSN). Accurate distance measurement between sensor nodes are used for localization in WSN and are very attractive for advanced wireless health-care applications. Several techniques such as received signal strength (RSS), angle of arrival (AoA) and time based approaches such as time-of-arrival (ToA) and time-difference-of-arrival (TDoA) exist for distance measurement between sensor nodes. ToA ranging approach is simple but require synchronization to give good precision. Clock synchronization between the transceivers is limiting the accuracy of the ToA estimation and hence increases the design challenges of the UWB ranging system [1]. For example, 1 GHz clock can give rise to 30 cm precision with high design complexity, which is not sufficient. For ranging accuracy better than 1 cm, we need a clock exceeding 30 GHz, which is quite challenging with the current CMOS process technology. In addition, multipath components and NLOS propagation adds to the complexity of the UWB ranging system.

Several reported systems for moderate data rates, low power and low complexity short-range communications and ranging may be found in the literature. In [2] a single-bit quantizer and auto-correlation followed by integration for pulse recovery is proposed. The actual time-of-arrival information is lost with the integration and limits the ranging accuracy to approximately 30 cm. The accuracy is <15 cm in [3] demanding high speed ADCs. Our goal is to take the benefit of the wideband characteristics of impulse radio systems and seek a system solution which dramatically cuts down the implementation cost while exploring the inherent capability of fine-timing

resolution. Using a two-way ranging between the sensor nodes, the distance can be estimated by measuring the symbol round-trip time without a common reference clock [4].

In this paper, we present a silicon implementation of a clockless ranging system using a continuous-time symbol detector based on CTBV technique. This system is also *coherent* preserving the time-of-arrival information throughout the processing and by improving the sampler, more accurate time-of-arrival measurements are feasible. It is possible to realize such an IR-UWB system using CTBV signal coding, which was proposed in [5]. The CTBV technique is fundamentally a new approach to signal processing simply based on inherent, CMOS process-dependent gate delays combined with a coarse quantizer or single-bit ADC. In CTBV domain, the signal value is binary but time is kept continuous, which combines the speed of analog signal processing with some of the advanced processing capabilities of digital signal processing. An important property of the simple CTBV bit-stream is the accurate sampling time in principle eliminating quantization errors.

Symbol detectors using running cross-correlation is normally power consuming, but CTBV coding enables power-efficient CMOS solutions doing template matching [6]. Calibration procedures must be adapted for adequate timing precision. In addition information coding as short-duration pulses must be maintained throughout the processing chain. Hence continuous-time storage elements like delayline must be carefully designed such that process variation affecting pulse width and chip interval of the incoming symbol are reduced. The CTBV symbol detector with careful design of the delay elements combined with proper initial calibration enables high precision ranging. The finite time resolution of the clock is overcome by CTBV coding and the ranging accuracy is not limited by the clock resolution.

In a master-slave configuration, the symbol time-of-flight is a combination of transceiver processing time and actual time-of-flight of the transmitted symbol. The fundamental idea is to keep all processing times constant leaving variations to only distance changes between the transceivers. The CTBV ranging system is very attractive for the emerging medical field such as wireless capsule endoscope, which requires high accuracy localization and in-body communication.

The rest of the paper is organized as follows. II gives an overview and principle of the presented ranging system. The symbol detector block is explained in III. Measured results are

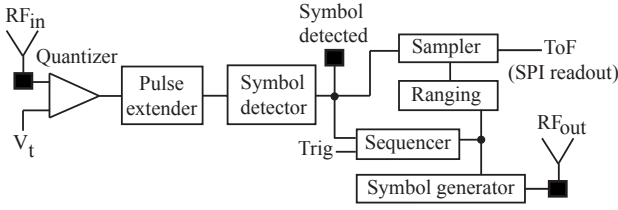


Fig. 1. CTBV ranging system.

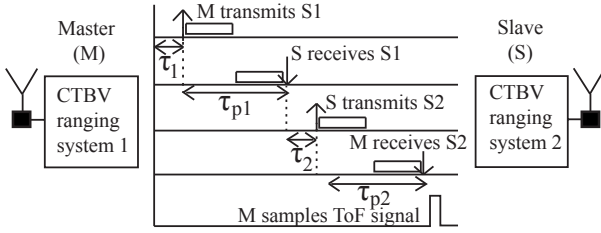


Fig. 2. Two-way ranging scheme to measure ToF.

detailed in IV. V concludes this work.

II. CTBV RANGING SYSTEM

Ranging in CTBV system is implemented using a continuous-time full RAKE symbol detector in a single-chip CMOS transceiver as shown in Fig. 1. The key blocks of the ranging transceiver are a quantizer, a continuous-time symbol detector, a symbol generator, sequencer and a ranging and thresholding sampler doing the actual ToF measurement. A single ended 1-bit quantizer with tunable threshold setting is quantizing the incoming RF signal to binary values. The quantizer is essentially a comparator continuously comparing the incoming signal to a threshold voltage, giving a digital output which is a CTBV coded signal [7]. The sampler is based on digital building blocks which are not clocked, but rather uses continuous-time signal processing and are suitable for modern CMOS technology with only two quantization levels. The key element of the sampler is a delayline based clock-signal for sampling the input bitstream with high sampling rate. An implementation in 90 nm CMOS technology has measured a sampling rate equivalent to 23 GHz [8] and can increase further with better technology. The sequencer is used for symbol transmission. Accurate chip interval of the PN-coded symbols are ensured by calibration. The sequencer can be programmed to generate the required pseudo-noise (PN) symbol pattern. During master mode, sequencer can be triggered to start the communication and in slave mode immediate transmission upon symbol reception is guaranteed. The sequencer and the symbol generator together forms the transmitter. More detailed description of this transmitter is beyond the scope of this paper and will be published in future.

The basic principle to measure symbol round trip time using the CTBV ranging system is explained in Fig. 2. In a master-slave configuration, master is triggered to initiate a communication and the sequencer transmits symbol S1 and

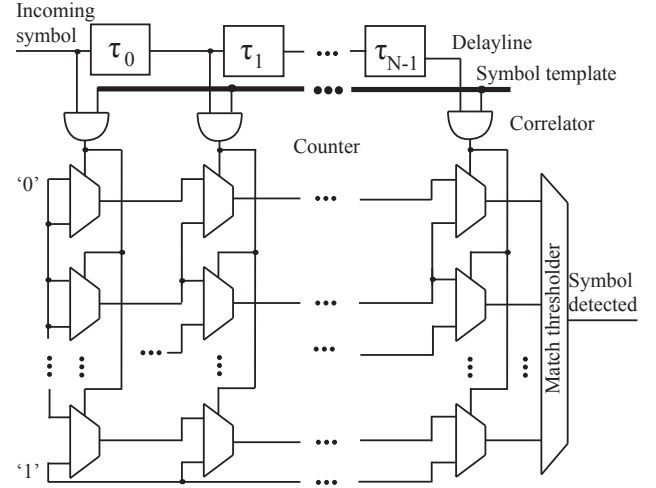


Fig. 3. Continuous-time symbol detector.

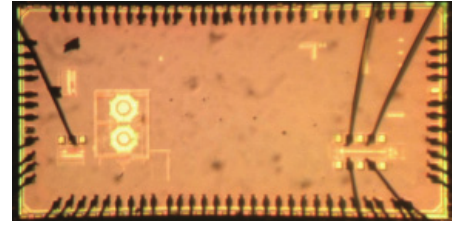


Fig. 4. Chip photograph of the CTBV ranging system.

starts the sampler. Initial delay (τ_1) corresponds to the start-up time of the symbol generator, which is almost negligible. The slave receives the symbol S1 after some propagation delay τ_{P1} . The propagation delay (τ_{P1}) consists of the complete symbol transmission time in addition to the actual time-of-flight. The slave is continuously looking for the expected symbol (S1) using cross-correlation. After receiving S1, slave will immediately transmit symbol (S2). A constant processing delay (τ_2) is added. Now master receives symbol S2 after (τ_{P2}) delay and the detection time of incoming symbol is sampled in continuous time by the sampler. Based on the symbol round-trip time, the ToF can be estimated precisely. There will be some jitter and small variations in the system, which may impact the resolution accuracy of the ranging system, but these errors may be reduced both by calibration and proper circuit design.

III. CONTINUOUS-TIME SYMBOL DETECTOR

Fig. 3 shows the block diagram of the continuous-time symbol detector, which performs running cross-correlation with a reference template operating in the CTBV domain. To limit the hardware complexity, the symbol length was set to 32-bits, but is scalable to longer symbols for improved sensitivity. The delayline stores the incoming bitstream. For continuous-time system, we will sample the pulse only during the pulse width and hence sufficient pulse width need to be maintained across the delayline. For proper operation, bit delays are calibrated

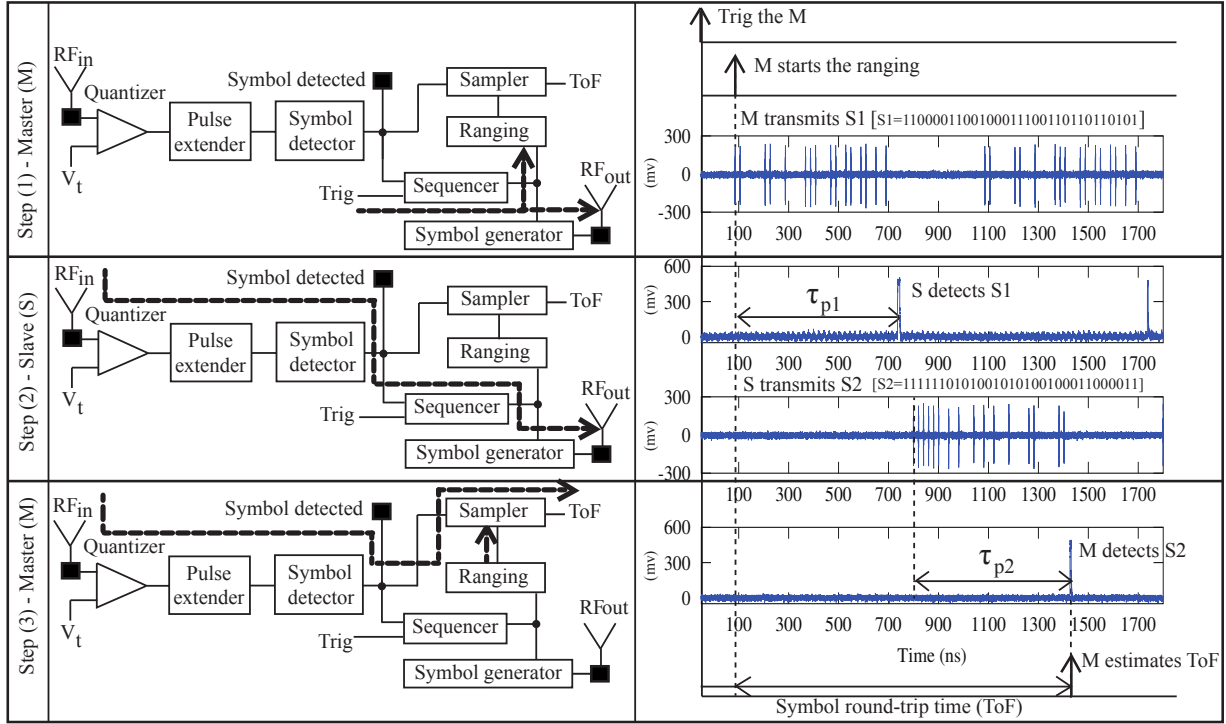


Fig. 5. ToF method using CTBV ranging system. Left: Signal flow path in master & slave transceivers. Right: Corresponding measured results.

with approximately 50 ps resolution. The tuning or calibration procedure is done by programming each delay element to the required chip interval by coarse-tune and fine-tune register settings using a SPI interface to the microcontroller. The symbol detector uses a digital correlator based on AND gate, which continuously correlates the incoming symbol and detects the number of pulses (bit '1') in the received symbol. By avoiding fast clocking registers and parallelism, the correlator is very power-efficient and at the same time performs high speed correlation without any hardware overhead. The detected number of pulses at the correlator output are continuously measured as a thermometer-coded digital value. The expected number of correlation matches is given by the number of bits set in the symbol. Using the match threshold, a statistical estimation of received pulses in the incoming symbol bitstream can be performed.

IV. MEASURED RESULTS

The proposed CTBV ranging system was designed and fabricated using the TSMC 90 nm low power process technology. The chip photograph of the CTBV ranging system is shown in Fig. 4. The measurement was conducted in an unshielded laboratory environment. The actual procedure with signal flow path for ToF measurement and its corresponding measured results are verified in 3 steps as shown in Fig. 5. (1) Master transmits a 32-bit symbol (S1) with a chip interval of 20 ns and starts the ranging circuit. (2) A digital pulse is seen at the slave output when it receives S1. Slave transmits symbol (S2) after a short processing delay. (3) Once master detects

S2, the sampler starts sampling the symbol detected pulse. The ranging circuit is a delay block, which is continuously running to capture the symbol detected pulse in the sampler. The rise time of the symbol detected pulse in the sampler gives the measure of total symbol round-trip time. Sampler consists of a delayline, with 64 taps and 64 samplers (D flip-flops), each connected to a 16-bit counter. When the signal arrives, each of the taps on the delayline triggers the corresponding sampler and the counter is incremented. The unit delay in the sampler is around 100 ps, which limits the ranging accuracy in the current implementation but improves with technology scaling.

The distance between the transceivers can be estimated using ToF measurement, ($d = c \cdot \frac{T_{ToF}}{2}$), where c is the speed of light. T_{ToF} ($\approx T_{rise} \cdot 100$ ps) is directly measured from the sampler output because all other logic delays ($\tau_1, \tau_{P1}, \tau_2, \tau_{P2}$) in both master and slave are almost constant during each communication. For the ranging experiment, the relative distance between master and the slave PCB board are varied in step of 1 cm each and the corresponding distance is measured using the ToF method as shown in Fig. 6. The standard deviation of the residuals after linear fitting, i.e. the ranging accuracy of the proposed system, is around 1.4 cm.

For each distance separation between the transceivers, the master is triggered 5000 times and the corresponding sampler output is plotted as shown in Fig. 7(b). The rise time of the sampler output in the x-axis, i.e. the ToF is varying with the distance. The sampler count in the y-axis tells us the

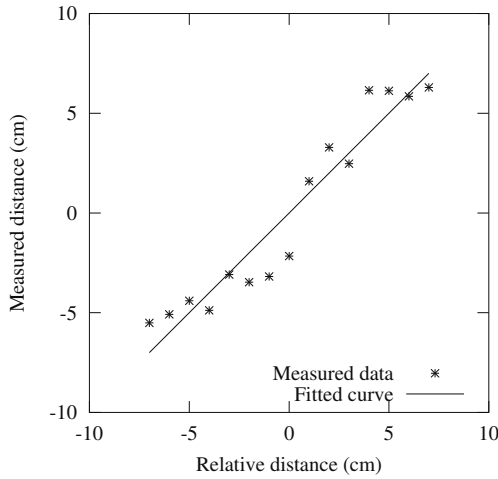


Fig. 6. Measured distance using ToF method versus actual distance (Fitted curve is calibrated against distance offset and sampling rate).

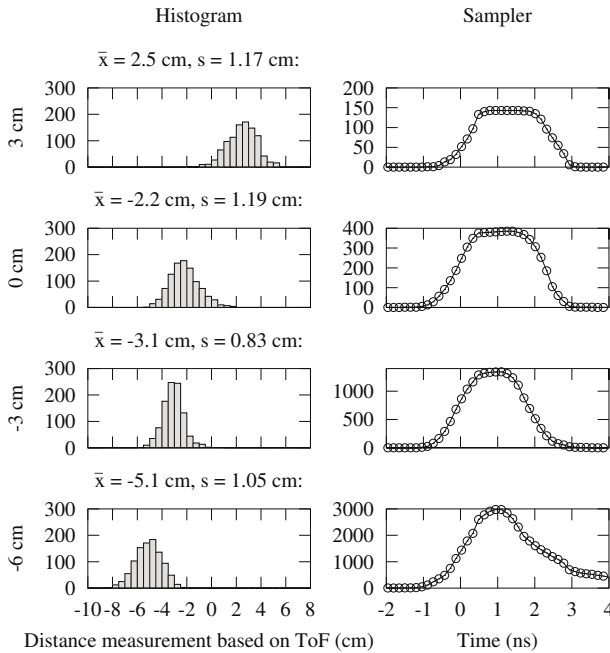


Fig. 7. (a) Measured distance. (b) Sampler output (ToF).

probability of a successful symbol detection. As the distance is increased, the probability of a successful symbol detection is reduced, which is visible in the plot. The data in the plots are based on 1000 repeated ToF measurements. By repeating the measurements, the processing gain is improved. The histogram plot for the corresponding distance between the transceivers is shown in Fig. 7(a). The jump in the plot at 3 cm is due to the varying threshold on the incoming signal as a function of distance. The quantizer threshold setting is a trade-off between false and missed symbol detections. The threshold setting can be challenging due to noise which impacts the performance of the system. We are working on extended measurements with

TABLE I
SUMMARY AND COMPARISON OF CTBV RANGING SYSTEM

Performance	This work	[2]	[3]
Process	90 nm	130 nm	180 nm
Area	3.6 mm ²	8 mm ²	17.2 mm ²
Supply Voltage	1.2 V	1.2 V	(1.8-1.2) V
Clockless detection	Yes	No	No
Bit-rate (Mbit/s)	1	31	0.25 to 20
Power	0.96 mW (digital) 4.6 mW (analog)	34 mW (analog only)	5.3 nJ/bit (excluding ADC)
Ranging accuracy	≈ 1.4 cm	< 30 cm	< 15 cm

different threshold settings as a function of distance.

Summary of the current work and comparison with respect to reported works are shown in table I. The proposed ranging solution is adding ToF measurements to IR-UWB communication with better ranging accuracy. The current work is also low power with minimum area and hence suitable for a wireless capsule endoscope application.

V. CONCLUSION

The reported IR-UWB system with embedded ranging has been implemented in 90 nm CMOS technology and presented with measured results. Proper initial calibration compensate for process variations in symbol detector and provide precision for a continuous-time system. Preliminary chip results shows that the IR-UWB ranging system is possible for accurate ToF measurements with ranging accuracy (≈ 1.4 cm). Ranging resolution is not limited by the receiver complexity but limited by the available CMOS process technology. The IR-UWB ranging system is power and area efficient, giving superior ToF distance measurements, which can be used for wireless medical applications such as wireless capsule endoscope.

ACKNOWLEDGMENT

The research is being sponsored by Norwegian Research Council project number 187857/S10 and is carried out at the Nanoelectronics group, Department of Informatics, University of Oslo, Norway.

REFERENCES

- [1] S. Gezici *et al.*, "Localization via ultra-wideband radios: a look at positioning aspects for future sensor networks," *IEEE Signal Processing Magazine*, 2005.
- [2] D. Lachartre *et al.*, "A 1.1nJ/b 802.15.4a-compliant fully integrated uwb transceiver in 0.13μm cmos," in *IEEE International Solid-State Circuits Conference*, Feb. 2009.
- [3] Y. J. Zheng *et al.*, "A 0.92/5.3nJ/b uwb impulse radio soc for communication and localization," in *IEEE International Solid-State Circuits Conference*, Feb. 2010.
- [4] L. Joon-Yong and R. A. Scholtz, "Ranging in a dense multipath environment using an uwb radio link," *IEEE Journal on Selected Areas in Communications*, 2002.
- [5] H. A. Hjortland and T. S. Lande, "Ctbv integrated impulse radio design for biomedical applications," *IEEE Biomedical Circuits & Systems*, 2009.
- [6] S. Sudalaiyandi *et al.*, "Power-efficient ctbv symbol detector for uwb applications," in *IEEE International Conference on Ultra-Wideband*, Sep. 2010.
- [7] T. A. Vu *et al.*, "Continuous-time cmos quantizer for ultra-wideband applications," in *IEEE International Symposium on Circuits & Systems*, Jun. 2010.
- [8] H. A. Hjortland, D. T. Wisland, T. S. Lande, C. Limbodal, and K. Meisal, "Thresholded samplers for uwb impulse radar," in *IEEE International Symposium on Circuits & Systems*, May 2007.

Paper 9 Live demonstration: IR-UWB transceiver with localization based on ToF technique suitable for wireless capsule endoscopy

S. Sudalaiyandi, H. A. Hjortland, T.-A. Vu, O. Naess, and T. S. Lande. “Live demonstration: IR-UWB transceiver with localization based on ToF technique suitable for wireless capsule endoscopy”. In: *Biomedical Circuits and Systems Conference (BioCAS), 2012 IEEE*, pp. 89–89, Hsinchu, Nov. 2012. doi:10.1109/BioCAS.2012.6418495. © 2012 IEEE. Reprinted with permission.

My Contributions to This Paper

This is a short paper for a live demo where time of flight measurement was demonstrated in real time. I helped create the chip which is used here, and I also helped create the measurement software.

Live Demonstration: IR-UWB Transceiver with Localization Based on ToF Technique Suitable for Wireless Capsule Endoscopy

Shanthi Sudalaiyandi, Håkon A. Hjortland, Tuan-Anh Vu, Øivind Næss, and Tor Sverre Lande
Dept. of Informatics, University of Oslo, Norway,
Email: shanthi@ifi.uio.no

Abstract—In this work, a unique continuous-time binary-value (CTBV) IR-UWB transceiver based on time-of-flight (ToF) technique is demonstrated. The demonstration includes real-time communication between master and slave nodes and its corresponding ToF estimation through live plotting which aids in localization. The CTBV ranging system is adding both localization and communication in a single chip (3.6 mm^2) to provide precision ranging (ranging accuracy $\approx 0.9 \text{ cm}$) with minimal area suitable for wireless capsule endoscope localization [1][2].

I. INTRODUCTION

Wireless capsule endoscopy (WCE) for gastrointestinal tract is an emerging medical field which can benefit from IR-UWB technology due to low power, small size, and cost efficient simple structures. However accurate capsule localization is still challenging due to significant signal attenuation as well as varying propagation speed. Along with the above, capsule must be small and power budgets are strongly limited. By incorporating the capsule in a wireless sensor network (WSN), trilateration may be used for accurate localization. Accurate capsule localization can be achieved provided distance measurements of EM-wave propagation between external WSN node and the capsule is accurate. For accurate EM-wave distance measurements, popular solutions based on received signal strength (RSSI) does not provide sufficient precision. Although several alternatives have been proposed, time-of-flight (ToF) measurements are usually assumed to give superior precision. But the clock synchronization between the transceivers and the clock frequency limits the ToF accuracy.

In this paper, a single chip solution including both communication and localization suitable for WCE is demonstrated. The transceiver is implemented using continuous-time binary-value (CTBV) coding [3] and localization based on time-of-flight (ToF) technique without a common synchronizing clock. Exploring the CTBV coding scheme, we are able to circumvent synchronization issues as well as maintain ranging precision without clock limitation on the ToF accuracy.

II. DEMONSTRATION SETUP

The demonstration setup consists of a CTBV ranging chip (both master (M) and slave nodes (S)) mounted on a wooden rig as shown in Fig. 1(a). The equipment will include a laptop, CTBV ranging chip (transceiver), Olimex microcontroller board (SAM7-H256) and a wooden rig. The Olimex microcontroller board is used to set the required control and data settings

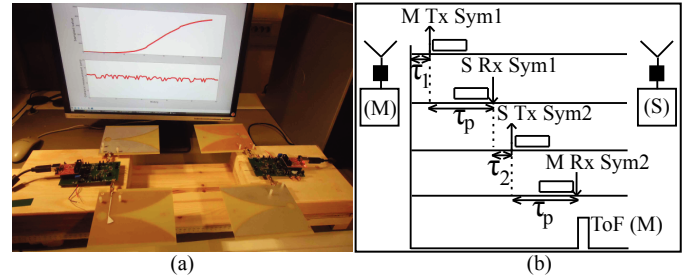


Fig. 1. (a) Demo setup. (b) Two-way ranging scheme to measure ToF.

through the SPI interface related to the transceiver operation. The SPI interface is connected to the laptop through the USB and programmed using Python language. The master and slave node is fixed on the wooden rig and the distance separation between them can be varied in precise centimeter increments. The principle of ToF estimation is depicted in Fig. 1(b). The demo is initiated by triggering master through SPI to start the symbol (Sym1) transmission and slave after receiving Sym1 issues immediate transmission (Sym2). Master receives Sym2 and the real-time ToF is measured using sampler. The measured ToF value is displayed on screen using Python live plotting.

III. VISITOR EXPERIENCE

The visitor can view the communication and localization of the CTBV ranging system on the laptop screen. The distance (in precise centimeter) between M and S nodes can be slowly varied by the visitor to view the live ToF and its distance estimation on screen. The visitor can also vary the input quantizer threshold and symbol detector match threshold to improve the sensitivity (symbol detections) of the ranging system. One need to stand in front of the wooden rig to avoid any interference which might impact the ToF measurement.

REFERENCES

- [1] S. Sudalaiyandi, H. A. Hjortland, O. Næss, T. A. Vu, T. S. Lande, and S. E. Hamran, "Ir-uwB transceiver with localization for wireless capsule endoscopy," in *To be Submitted to TBioCAS*.
- [2] S. Sudalaiyandi, H. A. Hjortland, T. A. Vu, O. Næss, and T. S. Lande, "Continuous-time high-precision ir-uwB ranging-transceiver in 90 nm cmos," in *Accepted in IEEE Asian Solid-State Circuits Conference*, 2012.
- [3] H. A. Hjortland and T. S. Lande, "Ctbv integrated impulse radio design for biomedical applications," *IEEE Biomedical Circuits & Systems*, 2009.

Paper 10 Continuous-time symbol detector for IR-UWB rake receiver in 90 nm CMOS

S. Sudalaiyandi, T.-A. Vu, H. A. Hjortland, O. Naess, and T. S. Lande. “Continuous-time symbol detector for IR-UWB rake receiver in 90 nm CMOS”. In: *Circuits and Systems (APCCAS), 2012 IEEE Asia Pacific Conference on*, pp. 487–490, Kaohsiung, Dec. 2012. doi:10.1109/APCCAS.2012.6419078. © 2012 IEEE. Reprinted with permission.

My Contributions to This Paper

This paper elaborates on the continuous-time symbol detector. I helped design the time-of-flight system and the circuits it consists of, including the continuous-time symbol detector. I helped with the improvements to the delayline design, with the design of the pulse width controller, and with the calibration algorithm used. I also helped create some of the plots.

Continuous-Time Symbol Detector for IR-UWB Rake Receiver in 90 nm CMOS

Shanthi Sudalaiyandi, Tuan-Anh Vu, Håkon A. Hjørtland,
Øivind Næss and Tor Sverre Lande

Dept. of Informatics, University of Oslo, Norway, Email: shanthi@ifi.uio.no

Abstract—In this paper, a coherent continuous-time symbol detector suitable for a simple RAKE receiver using single-bit correlation is reported. A working chip with fully integrated digital symbol detection without any clocks, ADC and filters for simple and power efficient CMOS implementation has been presented. Cross-correlating RAKE receiver with on-off keying is the most challenging receiver due to its hardware implementation complexity. Also process variations in 90 nm CMOS technology impose significant design challenges to the delayline. Optimal design of the delay elements and the use of pulse width controller blocks (PWC) along the delayline allows narrow pulses through the system significantly reducing the process variation impact. The optimized delayline combined with the high-speed counter and the match thresholder constitute a complete continuous-time IR-UWB RAKE receiver. With a data rate of 1.5 Mbit/s and a power consumption of 2.36 mW, operation of the symbol detector in continuous-time is verified with measured results. The symbol detector along with the ranging system in a master-slave configuration can estimate time-of-flight (ToF) value which aids localization.

I. INTRODUCTION

Many published works shows the great potential of ultra-wideband impulse radio (IR-UWB) towards short-range communications requiring low power and low data-rate [1]. Channel estimation, accurate synchronization and high resolution clocks forms the bottleneck to realize a power efficient UWB solutions [2]. UWB signals require RAKE or correlating receivers to take full advantage of the available signal energy over a large bandwidth. Because of complexity constraints, a RAKE receiver processes only a subset of the total received multipath components. To mitigate the issues of RAKE receivers, different configurations have been proposed [3]. Most implementations sacrifice performance for low complexity operation and lacks a power efficient CMOS implementation aiming towards a cost efficient single chip solution. Even if the receiver complexity is overcome by digital techniques, current solution requires ADC with high sampling frequency to sample and quantize the received signal [3].

A different approach for IR-UWB radio named Continuous-Time Binary-Value (CTBV) signal processing have been proposed by [4]. The CTBV approach is simply based on inherent CMOS process-dependent gate delays combined with a coarse quantizer. By avoiding clocks, power efficient solutions are feasible in standard CMOS while the performance of the CTBV solutions are only limited by the gate delays and scales well with process technology. A similar approach is proposed in [5] named Continuous-Time Discrete Amplitude

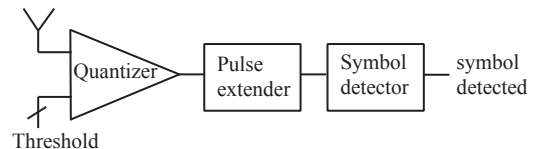


Fig. 1. Block diagram of the IR-UWB rake receiver.

(CTDA). The continuous time sampling systems compared to the discrete time avoids quantization noise and clocks. Also continuous time systems has fewer distortion components at frequencies within the band, input activity dependent power dissipation and fast response to sudden input variations [6].

A CTBV symbol detector using time-domain implementation can provide power-efficient solutions in conventional correlator receivers with minor increase in receiver complexity. The symbol detector in [7] is based on cross-correlation, storing the incoming signal in a suitable delayline (matching to the chip interval) and continuously performing the cross-correlation of the received symbol with the reference template. The measurement results showed around 1 ns of pulse shrinking per delay element (DE) throughout the delayline. Given the fact that the IR-UWB pulses can have components less than 1 ns, the incoming pulses of the symbol may disappear in the delayline. Pulse shrinking in the delayline posed a significant challenge to realize a CTBV symbol detector on chip. Also since CTBV system does not use any clocks for symbol detection, process variation and mismatch are very problematic leading to synchronization issues.

In this paper, we present a CTBV based symbol detector suitable for continuous-time IR-UWB RAKE receiver. We have shown that with careful design of the delay elements (DE) combined with proper initial calibration have reduced the process variation impact on the delayline. Also we have implemented pulse width controllers (PWC) in the delayline maintaining sufficient pulse widths throughout the delayline.

The rest of the paper is organized as follows. II gives an overview of the IR-UWB RAKE receiver. III gives the detailed description of the symbol detector and the design methodology to mitigate the issue of pulse shrinking. Measured results of the chip are explained in IV. V concludes this work.

II. IR-UWB RAKE RECEIVER

Fig. 1 shows the block diagram of a IR-UWB rake receiver in continuous-time. A single ended quantizer with tunable

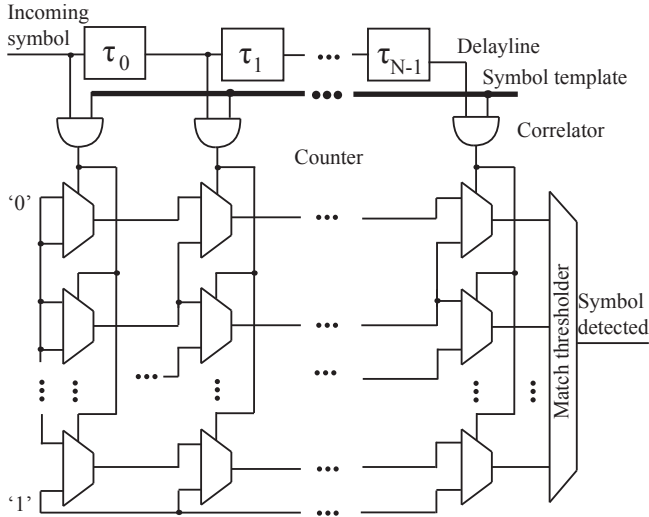


Fig. 2. Schematic of the symbol detector.

threshold setting is at the receiver front-end. The quantizer (1-bit) is essentially a thresholder continuously comparing the incoming signal to a threshold voltage, giving a single-bit digital output which is a CTBV coded signal [8]. The optimal threshold value of the quantizer is a trade-off between noise and the narrow pulse width signals. Best solution is to set the threshold high enough to get rid most of the noise and other ringing effects. Using pulse extender, the signal is reconstructed with sufficient pulse width to pass through the deep digital logic in the symbol detector. The pulse extended CTBV signal is then cross-correlated with the symbol template in the symbol detector block. Thus performing both threshold detection and the correlation detection on the incoming symbol, we are able to detect the symbol.

III. SYMBOL DETECTOR ARCHITECTURE

In this paper, the focus is on the continuous-time symbol detector architecture and Fig. 2 shows the schematic of the symbol detector. The quantized output may be very short pulses and hence pulse extended using the tunable pulse extender. The basic working principle of the continuous-time symbol detector is as shown in Fig. 3. The incoming pulse sequence is stored in a delayline and cross correlated with the input template in continuous-time. The symbol detector is based on on-off keying (OOK) and hence correlation refers to the matching of incoming '1's with respect to the stored reference template. The correlated output are counted and coded as a thermometer coded digital value using the high speed counter. The counter output may then be thresholded with an appropriate detection level using the match thresholder. By setting different values in the match thresholder, we can statistically detect the number of pulses present in the incoming symbol.

A. Delayline

The delayline consists of 31 delay elements (DE) to process the incoming 32-chip symbol. Each delay element is

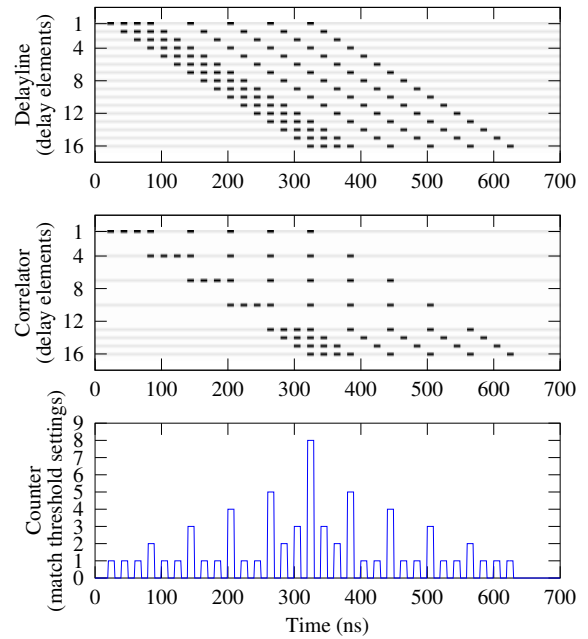


Fig. 3. Principle of symbol detection (symbol pattern-1111001001001001).

implemented using a coarse and fine tune in cascade, wherein coarse delay element is designed using cascaded NMOS and PMOS structures and fine delay element is designed using the minimum sized inverters (limited by the process technology). By tapping the delayline in suitable intervals matching the chip interval of the transmitted symbol, the symbol will appear simultaneously as a bit-pattern. The process dependent variation, mismatch, and jitter can significantly impact the delayline, which causes variation both in chip interval and pulse width of the incoming symbol.

To overcome the chip interval variation, each delay element is calibrated to match the chip interval of the incoming symbol. The calibration is an initial setup which can be done without the need for any additional hardware. The tuning or calibration procedure is achieved by programming each delay element through coarse-tune and fine-tune register settings using a SPI microcontroller. With respect to the first delay element, the second delay element is tuned such that they have maximum pulse overlap between them. This can be verified at the symbol detector output by setting the match thresholder settings to 2. A pulse will be detected at the output if the above two delay elements are calibrated properly in time. Similarly all other delay elements are also tuned in time with respect to the first delay element.

The pulse width variations in the delayline are significantly reduced by optimal design technique and also the use of pulse width controller blocks (PWC) in the delayline. The delayline consists of long chain of multiplexers and inverters, which forms asymmetrical capacitive loading to the signal path. Hence the inverters in the delay element are carefully designed such that their pulse width variation is constant irrespective

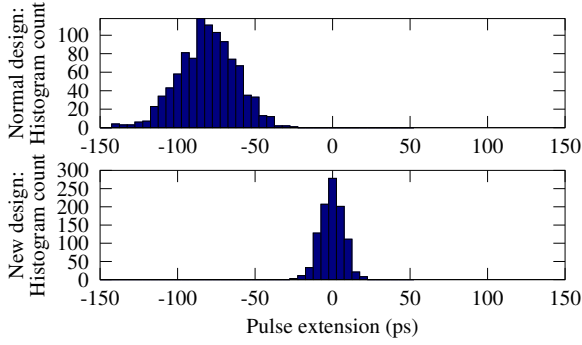


Fig. 4. Monte Carlo simulation results showing pulse width variation in the fine delay element.

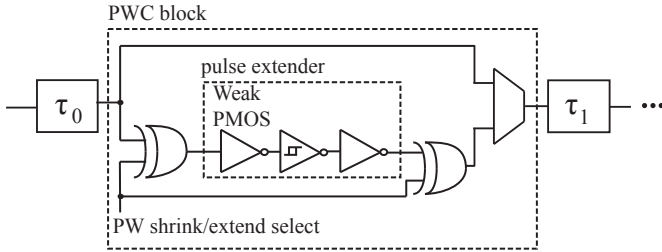


Fig. 5. Schematic of the PWC block along the delayline.

of different capacitive loading. Thus any shrinking/extending the pulse width is balanced out by the inverting and non-inverting logic present in the delayline. Fig. 4 shows the post-layout Monte Carlo simulation of fine delay element (8 buffers with mux8) with optimal designed inverters and the standard deviation is around 7 ps. Around 124 of these fine delay elements are in cascade in the delayline and the inverter with optimal sizing has improved pulse width variation compared to the normal inverter design.

Further variation in the pulse width can also be tuned by the pulse width controller blocks (PWC), which is used in the delayline as shown in Fig. 5. The pulse width can be shrunk or extended by a pulse extender, which uses a weak PMOS to vary the rising edge of the pulse. As variation and mismatch causes an uncorrelated effect on the delayline, the pulse width of the incoming signal after each delay element is either extended or shrunk using the PWC block.

B. High Speed Correlation

There are different analog and digital correlators available that are suitable for a wideband correlation. Straight forward implementations of the digital correlators are very power hungry demanding fast DSPs to perform several giga multiplications and adds per second [9]. On the other hand, analog implementations are very sensitive to noise and process variations [10]. In the current work, the symbol detector uses a digital correlator based on AND gate, which continuously correlates the incoming symbol and detects the number of pulses (number of correct matches) in the received symbol. By avoiding fast clocking registers and parallelism, the correlators are very power-efficient and at the same time performs high

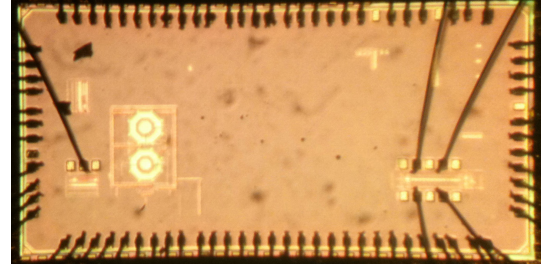


Fig. 6. Chip microphotograph.

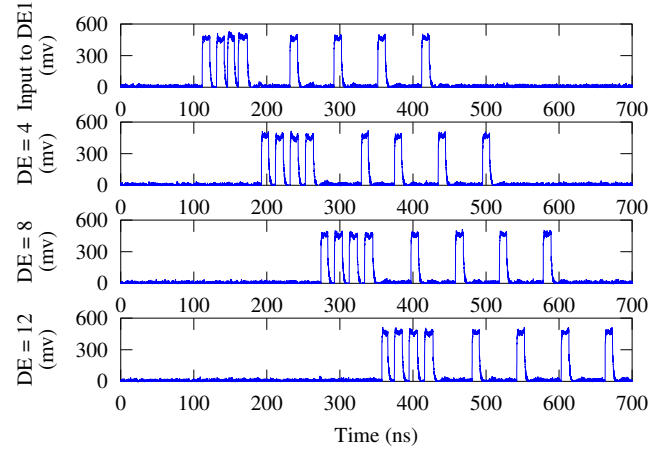


Fig. 7. Measured correlator output and its delayed versions across the delay elements (DE) corresponding to the symbol pattern-1111001001001001.

speed correlation without any overhead. The detected number of pulses at the correlator output are continuously routed to the counter output as a thermometer-coded digital value, which simplifies the readout circuitry.

C. Match Thresholder (MT)

The continuous-time counter output is analyzed by the match thresholder block as shown in Fig. 2. The match thresholder consists of a multiplexer, which can be programmed through the SPI microcontroller. Depending on the number of pulses detected in the incoming symbol, the appropriate threshold can be set in the match thresholder for a statistical symbol detection estimate.

IV. MEASURED RESULTS

A continuous-time symbol detector with 32-chips was designed and fabricated using TSMC 90 nm low power process technology and Fig. 6 shows the chip microphotograph of the CTBV ranging system along with the symbol detector.

To verify the symbol detector chip, a symbol pattern was transmitted using the symbol generator in the ranging system and received by the symbol detector in continuous-time. The correlator output across the delay elements (DE) are shown in Fig. 7. Each of the delay elements are tuned to match the chip interval (20 ns) and hence the output at DE = 12 is around 240 ns delayed, with an initial delay corresponding to the symbol transmission. With optimal design of delay

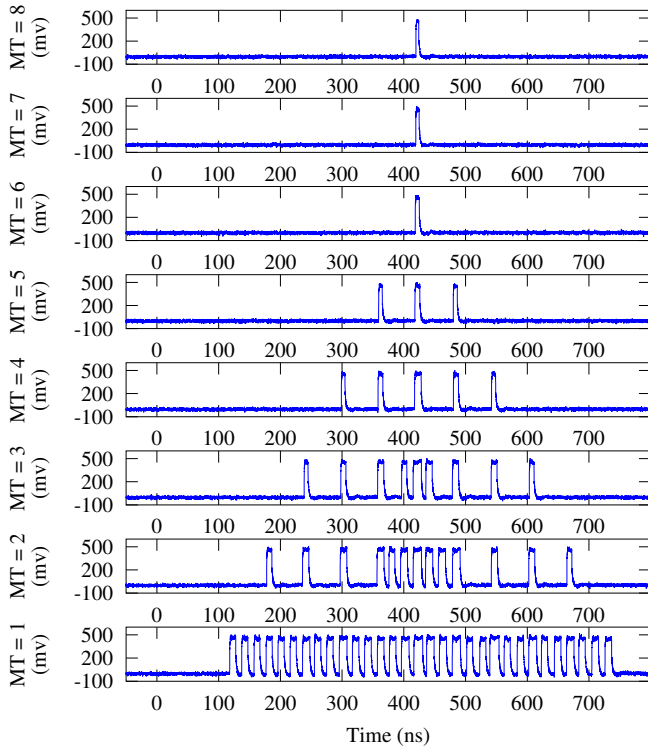


Fig. 8. Measured symbol detection for different match threshold (MT) settings corresponding to the symbol pattern-1111001001001.

elements and the initial calibration technique along with the use of PWC blocks, we are able to see the pulses across each delay element without getting shrunk. With different threshold settings enabled in the match thresholder, a statistical estimation of the received symbol can be accomplished by the symbol detector as shown in Fig. 8. Since the incoming symbol has 8 pulses, the counter output (thermometer coded digital value) can be selected up to 8 through the match thresholder for the symbol detection. The impact of jitter can also be reduced to some extent by tuning the delay element such that the output depends on the less jittery delay element (nearest delay element) than the farthest delay element.

The clockless and continuous-time symbol detection feature presented in this paper is highly appreciated for precision ranging in localization. The IR-UWB CTBV ranging system using the symbol detector in a master-slave configuration can be used to estimate the time-of-flight (ToF) without any clock synchronization. If the incoming symbol matches the symbol template and with appropriate match threshold settings, the symbol pattern is detected. Immediately, the slave transmits another symbol pattern to master and when master receives this symbol, the symbol detector preserves the symbol round-trip (ToF) timing information which aids localization. The coherency of the symbol detection is obtained since cross-correlation is running in continuous-time and the detection pulse will be issued immediately when matching occurs.

The performance summary including ToF measure with different chip intervals are as shown in table I. The data rate

TABLE I
PERFORMANCE SUMMARY OF SYMBOL DETECTOR

Chip interval (ns)	ToF (μ s)	Data rate (Mbit/s)	Power (mW)
18	1.322	1.65	2.37 (digital), 4.29 (analog)
20	1.326	1.5	2.36 (digital), 4.56 (analog)
22	1.343	1.37	2.36 (digital), 4.58 (analog)

is calculated as, $(1/((T_{chip} * N_s + T_{quiet}) * n_{repetition}))$ where T_{chip} is the chip interval and T_{quiet} (≈ 30 ns) is the assumed quiet period after every symbol transmission. $n_{repetition}$ gives the number of times symbol is transmitted and improves the processing gain of the system while reducing the data rate. N_s corresponds to number of chips, which can be increased further with increase in receiver complexity.

V. CONCLUSION

A coherent continuous-time symbol detector using running cross-correlation without any clocks at a data rate of 1.5 Mbit/s was designed in 90 nm CMOS technology and measured chip results are provided. The proposed novel symbol detector being power-efficient is a key block for IR-UWB receiver architecture, working towards a promising solution solving some of the fundamental implementation limitations for high precision ranging architectures.

ACKNOWLEDGMENT

The research is being sponsored by Norwegian Research Council project number 187857/S10 and is carried out at the Nanoelectronics group, Department of Informatics, University of Oslo, Norway.

REFERENCES

- [1] Y. J. Zheng, S. Diao, C. Ang, Y. Gao, F. Choong, Z. Chen, X. Liu, Y. Wang, X. Yuan, and C. H. Heng, "A 0.92/5.3nJ/b uwb impulse radio soc for communication and localization," in *IEEE International Solid-State Circuits Conference*, Feb. 2010, pp. 230–231.
- [2] W. Yiyin, L. Geert, and A. van der Veen, "Digital receiver design for transmitted reference ultra-wideband systems," *EURASIP Journal on Wireless Communications and Networking*, 2009.
- [3] Y. Chao and R. A. Scholtz, "Ultra-wideband transmitted reference systems," *IEEE Transactions on Vehicular Technology*, Sep. 2005.
- [4] H. A. Hjortland and T. S. Lande, "Ctbn integrated impulse radio design for biomedical applications," *IEEE Transactions on Biomedical Circuits and Systems*, pp. 79–88, Apr. 2009.
- [5] Y. W. Li, K. L. Shepard, and Y. P. Tsividis, "A continuous-time programmable digital fir filter," in *IEEE Custom Integrated Circuits Conference*, Sep. 2005, pp. 695–698.
- [6] B. Schell and Y. Tsividis, "A continuous-time adc/dsp/dac system with no clock and with activity-dependent power dissipation," *IEEE Journal of Solid-State Circuits*, pp. 2472–2481, Nov. 2008.
- [7] S. Sudalaiyandi, M. Z. Dooghabadi, T.-A. Vu, H. A. Hjortland, O. Naess, T. S. Lande, and S. E. Hamran, "Power-efficient ctbn symbol detector for uwb applications," in *IEEE International Conference on Ultra-Wideband*, Sep. 2010, pp. 1–4.
- [8] T. A. Vu, S. Sudalaiyandi, M. Z. Dooghabadi, H. A. Hjortland, O. Naess, T. S. Lande, and S. E. Hamran, "Continuous-time cmos quantizer for ultra-wideband applications," in *IEEE International Symposium on Circuits and Systems*, Jun. 2010, pp. 3757–3760.
- [9] M. Verhelst, W. Vereecken, M. Steyaert, and W. Dehaene, "Architectures for low power ultra-wideband radio receivers in the 3.1-5ghz band for data rates < 10mbps," in *International Symposium on Low Power Electronics and Design*, Aug. 2004, pp. 280–285.
- [10] S. Sriram, K. Brown, and A. Dabak, "Low-power correlator architectures for wideband cdma code acquisition," in *Conference Record of the 33 Asilomar Conference on Signals, Systems, and Computers*, 1999.

Paper 11 A continuous-time IR-UWB RAKE receiver for coherent symbol detection

S. Sudalaiyandi, H. A. Hjortland, and T. S. Lande. “A continuous-time IR-UWB RAKE receiver for coherent symbol detection”. *Analog Integrated Circuits and Signal Processing*, Vol. 77, No. 1, pp. 17–27, Oct. 2013. doi:10.1007/s10470-013-0120-0. © 2013 Springer Science+Business Media. Reprinted with kind permission from Springer Science+Business Media.

My Contributions to This Paper

This paper elaborates on the continuous-time symbol detector presented in earlier papers. I came up with much of the system design, and I also helped design many of the subcircuits such as the continuous-time symbol detector, the continuous-time counter, the pulse width measurement circuit, the sampler, and the symbol generator. I helped with the development of the optimal delay element, and I also helped create the delayline calibration algorithm. I created the symbol set used in the paper, and I helped with the measurement design, the measurement algorithms, and the measurement software. My SPI and microcontroller solution was used to control the chip.

A continuous-time IR-UWB RAKE receiver for coherent symbol detection

Shanthi Sudalaiyandi · Håkon A. Hjortland ·
Tor Sverre Lande

Received: 15 February 2013 / Revised: 6 August 2013 / Accepted: 11 August 2013 / Published online: 29 August 2013
© Springer Science+Business Media New York 2013

Abstract The ultra-wide bandwidth released for unlicensed use by FCC a decade ago has initiated significant research efforts. The large ultra-wide bandwidth is attractive not only for increased data transfer speed but may also be exploited for added functionality like high-precision ranging in wireless sensor networks. RAKE based receivers are preferred for ultra-wideband (UWB) technology due to wide bandwidth. However, designing RAKE based correlating receivers remains quite challenging. Correlating receivers are also power consuming due to high-speed DSPs, ADC and matched filter. Timing synchronization is another issue associated with correlating receivers. In this paper a impulse radio ultra-wideband (IR-UWB) RAKE receiver is presented utilizing a continuous-time binary value coding scheme for power-efficiency and coherent symbol detection without the need for synchronization to achieve precise ranging using time-of-flight technique. A working prototype of the IR ranging transceiver which uses the IR-UWB RAKE receiver is presented with measured high-precision ranging towards 1.4 cm.

Keywords IR-UWB RAKE receiver · Continuous-time binary value (CTBV) coding · Correlating RAKE receiver · High-precision ranging

1 Introduction

With its promise of high data rate over short distances at very low power consumption, ultra-wideband (UWB) technology has been emerging as an attractive research

field ever since the release of 3.1–10.6 GHz unlicensed frequency band by Federal Communication Commission (FCC). UWB has been characterized with some unique properties such as broad spectrum, and multipath immunity which makes them interesting research areas such as wireless sensor networks, RFID tags, wireless health care, short distance communication and localization [1].

Among different UWB technologies, impulse-radio ultra-wideband (IR-UWB) technology has the best potential in terms of low complexity, low power, low cost and good performance compared to others such as direct sequence code division multiple access and multiband orthogonal frequency division multiplexing. Impulse radio (IR) is a form of spectrum spreading several pulses are transmitted per information symbol without carrier. These short pulses are typically on the order of nanosecond or sub-nanosecond leading to fine time resolution [2, 3]. The use of symbols with extremely narrow pulses with gigahertz bandwidths ensures that the multipath reflections can be resolved with different delays or ignored which greatly reduces the multipath fading effects. IR technique does not require carriers and hence does not require the baseband components such as modulators, demodulators and other intermediate frequency stages. This reduces the transceiver complexity, size, cost and overall power consumption. Because of the above characteristics, several reports indicate that the IR-UWB technology is a promising solution for low power short-distance communications and high precision localization in wireless sensor networks [4].

The transmission of IR-UWB symbols results in many resolvable multipath components and hence correlating RAKE receivers are the most viable options to exploit the time-diversity inherent in multipath components. The challenge is to come up with architectures able to detect the very short low-energy pulses with minimum power

S. Sudalaiyandi (✉) · H. A. Hjortland · T. S. Lande
Department of Informatics, University of Oslo, Oslo, Norway
e-mail: shanthi@ifi.uio.no

consumption. Correlating RAKE receiver collects these multipath components, decodes them independently, and sums them finally to reconstruct the original signal from the power-limited transmitted energy [5]. Coherent receivers such as correlating RAKE receivers consists of correlator banks to match the incoming symbol with the symbol template. The complexity of these receivers increases with the number of fingers while detecting the maximum number of resolvable paths. Since coherent receivers are complex and computational intensive, different receiver architectures such as non-coherent receivers, transmitted reference, analog correlation receivers and all-digital architecture are proposed to trade-off performance for complexity in many ways [6–8].

Non-coherent receivers may be simple to implement but has performance degradation and also suffers from noise interference. All-digital architecture are power consuming due to the running cross-correlation, ADC and matched filtering. The receiver needs an ADC with high-sampling frequency to sample and quantize the received analog signal which is difficult with the available CMOS technology. Even with simplified types of ADC architectures [9], the power consumption seem to be significant. Most reported receivers does not solve the associated problems with high speed ADC and matched filter block completely. Another solution is the analog correlation receivers doing the correlation operation in the analog domain which reduces the high-sampling rate of the ADC and hence reduces the power consumption [10].

Precise timing synchronization is still another problem to overcome with high sensitivity coherent receiver. The alignment between the received symbol and the symbol template for correlation should be within picoseconds range which is quite challenging to implement with the current CMOS process technology. Even though high-sampling rate of the ADCs can be reduced, coherent receivers still need complex timing synchronization to improve receiver sensitivity which in-turn cannot avoid clock generators both in transmitter and receiver blocks [11].

In this paper, a correlating RAKE receiver utilizing continuous-time binary value (CTBV) coding is presented. The RAKE receiver does coherent symbol detection suitable for short-range communication and localization. The IR-UWB CTBV ranging system was first presented in [12] and the current work is more detailed study of CTBV RAKE receiver block. The receiver explores single-bit quantization and clockless symbol detection using cross-correlation on the received symbol with the reference template in continuous-time. The thresholding operation is simply a comparator doing single-bit quantization often sufficient when signal-to-noise ratio is low (<1). The thresholded output as a bit-stream is then cross-correlated with the symbol template using digital gates to estimate the

symbol detection in the incoming symbol. The advantages of the CTBV RAKE receiver compared to the conventional RAKE receivers are in many ways,

- The symbol detector does running cross-correlation in continuous-time without the need for precise timing synchronization. Even in the absence of timing synchronization, the RAKE receiver does coherent symbol detection by preserving the timing information of the received symbol.
- The complex power hungry synchronization circuits are completely avoided because continuous-time signal processing does not need clocks and also avoids aliasing and quantization noise [13, 14].
- Very high sampling rate is possible without the need for precise timing synchronization using coarse sampler based on CTBV coding [15].
- The capability to tolerate errors in the incoming symbol and also to detect the exact position of the pulse in the symbol is unique. A statistical estimation of the symbol match in the receiver is feasible.
- Thresholding the quantizer using swept threshold technique enables variable sensitivity combined with processing gain.
- RAKE receiver adopting continuous-time signal processing may cover most of the multipath components. Thus the implementation complexity with respect to the number of fingers compared to the conventional RAKE receivers have been simplified without sacrificing performance.
- RAKE receiver based on on-off keying (OOK) scheme is well suited for low power and low complexity applications as documented in [16].

2 Continuous-time rake receiver

The key blocks in the continuous-time RAKE receiver include a quantizer, a pulse extender and a symbol detector. Figure 1 shows the block diagram of a CTBV RAKE receiver. A single ended quantizer with tunable threshold setting is at the receiver front-end. The quantized output is pulse extended using pulse extender circuit and then cross-correlated with a symbol template in CTBV domain using the symbol detector block.

Threshold detection requires that the received signal exceeds some threshold well above the noise level. On the other hand correlation detection compares the received signal with a reference template and provides an output proportional to the cross-correlation with the reference signal.

Advantages of analog thresholding through quantizer with tunable threshold settings in continuous-time: (1) Best

suited for receivers with low signal-to-noise ratio. (2) Optimal threshold value can be adopted according to signal strength. Advantages of cross-correlation in continuous-time using digital gates as multipliers: (1) No need for precise timing synchronization. (2) Without clocks, the system is very power-efficient. Thus performing threshold detection and correlation detection on the incoming symbol, we are able to successfully detect the symbol which is almost buried in the noise.

3 Quantizer

The quantizer is essentially a thresholder continuously comparing the incoming signal to a threshold voltage, giving a single-bit digital output which is a CTBV coded

signal and detailed description is in [17]. The 1-bit quantizer (ADC) is quite advantageous in terms of simplicity, high gain and variable threshold which are suitable for continuous-time RAKE receiver. The received signal at the quantizer input is compared with the threshold voltage (V_{th}) providing a controllable quantization voltage level for input signal thresholding. For improved noise performance, the threshold voltage is set by a current (I_{th}) and converted to a suitable threshold voltage level matching the built-in inverter threshold. The front-end of the quantizer consists of a DAC and a potentiometer to tune the threshold of the input stage. The quantizer threshold can be varied by varying the DAC values through the SPI microcontroller.

In the continuous-time RAKE receiver, threshold voltage setting is critical and must be adopted according to incoming signal strength and noise levels. Figure 2

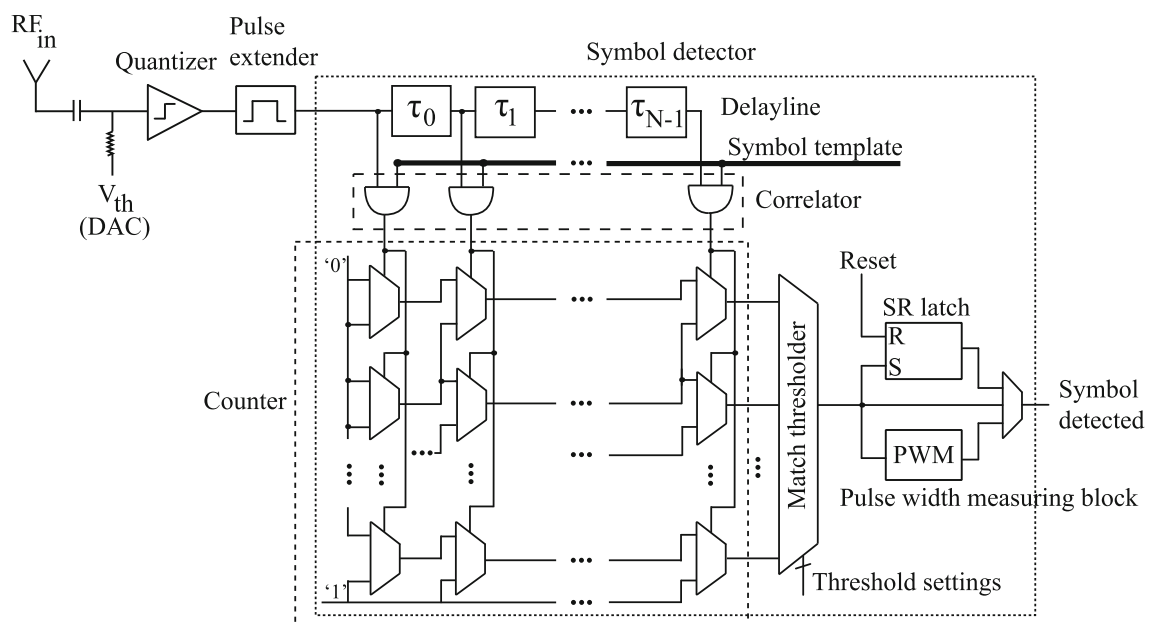


Fig. 1 Block diagram of the continuous-time RAKE receiver

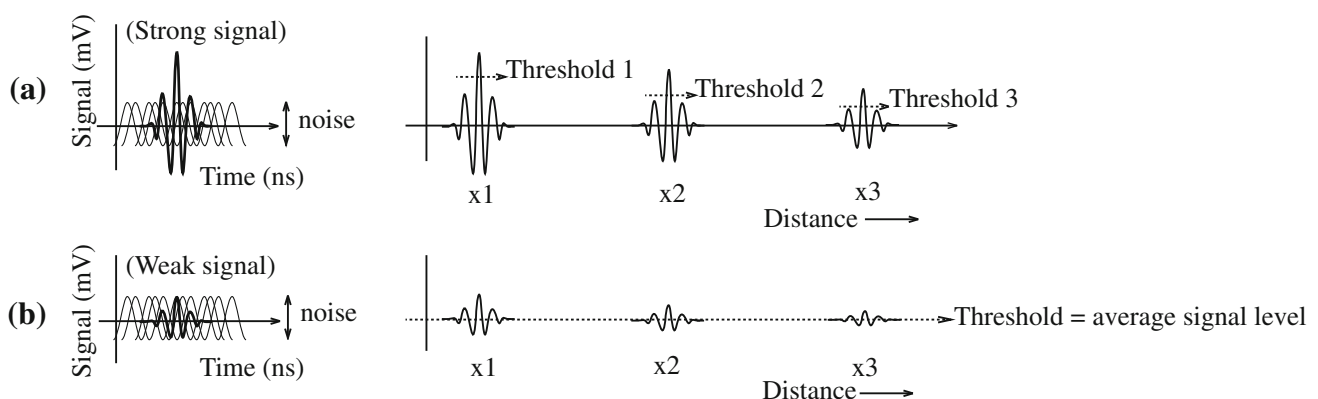


Fig. 2 Swept threshold technique and stochastic resonance

illustrates an example of a single pulse representation (for strong and weak signal) with respect to distance change between transmitter and receiver. Since quantizer threshold is a trade-off between noise and missed symbol detections, to find out the optimal threshold value is an interesting problem to be solved. In the continuous-time RAKE receiver, an optimal threshold value can be adapted using swept threshold technique or setting the quantizer threshold to the signal average value. The swept threshold technique for an impulse radar was already published in [18] and this work discuss the use of swept threshold technique for continuous-time RAKE receiver.

3.1 Swept threshold technique

In this paper, swept threshold technique is implemented where the optimal threshold value is adopted depending on distance change as illustrated in Fig. 2(a). The amplitude of the received signal is reducing with distance change between transmitter and receiver. Hence the quantizer threshold being a function of distance should also be swept with respect to distance change for optimal symbol detection. For example, as in Fig. 2(a), different threshold values such as threshold1, threshold2 and threshold3 are swept with respect to different distances x_1 , x_2 and x_3 . By sweeping the DAC value which changes the quantizer threshold, its corresponding symbol detections as a function of DAC value is plotted as shown in Fig. 3. If the threshold is too low, the output is always at low level. If the threshold hits the incoming signal, the quantizer output is a pulse with different pulse width depending on where the threshold hits the incoming signal. Conversely, if the threshold is too high (outside the signal), no symbols are detected and the output is at high level. This is an uninteresting case based on weak signals due to strong regulations. As a consequence we have to handle weak signals even buried in noise. Of course the better signal-to-noise ratio (SNR), the better detection quality. The clue here is that single-bit quantization is really good for low SNR detection when combined with processing gain. It still works for higher SNR (suprathreshold). However, once in the suprathreshold operation, increased threshold will not

improve symbol detection and may even lead to lost symbols for too high threshold voltages. The important observation is that stronger signals are detectable by increasing the threshold and weak signals are also detectable due to stochastic resonance. Depending on the required symbol detections and range, the DAC value is used which adapts the corresponding quantizer threshold.

The threshold can be set high such that less noise can pass through the system. With a high threshold, the received signal pulse width (PW) can vary depending on where the threshold will hit the incoming signal and hence can have shorter pulses into the system. With the help of pulse extenders, we can reconstruct the signal with sufficient PW to pass through the huge logic circuits. In case if multiple small pulses exist, pulse extenders in the incoming stages will absorb these small multiple pulses.

With respect to Fig. 2(b), if the signal is below the noise floor, the quantizer threshold can be set to the average signal level. We are aware that white noise is a random signal with a constant power spectral density. Now with the incoming signal, the white noise is modulated with respect to the incoming symbol pattern. With threshold set at the average signal value, and with white noise added, the pulses will sometimes exceed the threshold and the weak signal may be sensed. This behavior is called stochastic resonance. Stochastic resonance is a technique to enhance the SNR of the output with added noise [19]. For the continuous-time RAKE receiver, the input threshold may be set to signal average value of the quantizer so that weak signals are amplified above the noise floor and are detected. By fixing the input threshold to signal average, the input threshold is no more a function of distance change. However to figure out the average signal value is quite challenging given the complexity of the system and also noise makes it worse. The DAC value may be swept to vary the input threshold to figure out the average signal value of the quantizer. The average signal value can be determined if we get 50 % 1-bit and 50 % 0-bit because of the inherent property of white Gaussian noise. Although initial results indicate adequate threshold voltage settings may be found, additional research efforts are required for adaptive threshold voltage tuning.

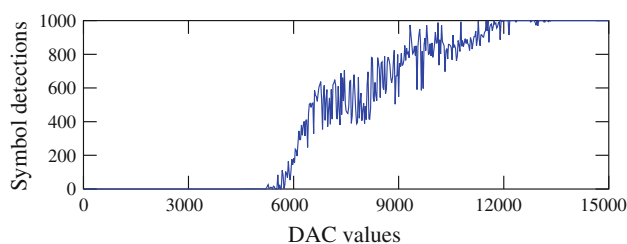
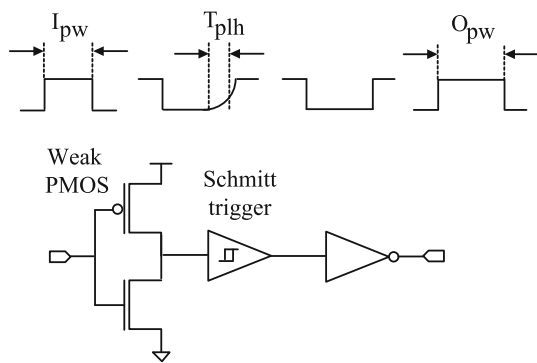


Fig. 3 Sweep of DAC values (quantizer threshold) versus symbol detections

4 Pulse extender

The quantizer output is very short digital pulses in the order of picoseconds. These short pulses are extended immediately with sufficient PW using the pulse extender so that they may be detected by digital logic in the symbol detector. In traditional clocked systems, a pulse is sampled within the clock period. On the other hand, in continuous-time system, a pulse is sampled only during the PW of that

**Fig. 4** Schematic of pulse extender

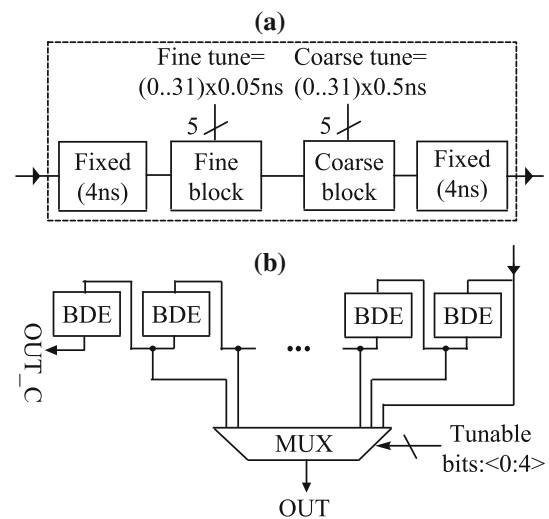
pulse. Hence the pulse need to be sufficiently wide during the sampling time. The pulse extender circuit is a fully digital schematic as shown in Fig. 4. The first inverter has asymmetric sizing of the gates to accurately control the pulse shape. The dimension of the PMOS and NMOS transistors are designed such that PMOS transistor is made weaker to prolong the rise time (T_{plh}) depending on the required extension. This mechanism is used to accurately control the PW extension. The pulse extender can be varied using the SPI interface such that (T_{plh}) is tuned according to the desired PW (O_{pw}) and is recovered by a schmitt trigger. The slow rise time is prone to jitter and process variations, but can be compensated by tuning, to some extent. The pulse extender can stretch the incoming signal to a maximum value of ≈ 12.8 ns with a resolution of 0.4 ns.

5 Symbol detector

The symbol detector is the key block in the continuous-time RAKE receiver which does the clockless coherent symbol detection in continuous-time. The symbol detector consists of delay line, correlator, a high-speed counter and a match thresholder [20]. The details of the symbol detector schematic is as shown in Fig. 1. The delay line in the symbol detector was designed with 31 delay taps to process the 32-chips symbol so that the implementation complexity is minimized.

5.1 Delay line

Delay line is one of the critical blocks in the continuous-time RAKE receiver. The delay line consists of $N - 1$ delay elements (DE) to process symbols of length N (pulse extended quantized CTBV signal). The cascaded DE stores the received signal matching the chip-length to store the history of the received signal. By tapping the delay line in suitable intervals matching the chip-rate of the transmitted symbol, the symbol code should appear simultaneously as a

**Fig. 5** (a) Schematic of delay element. (b) Implementation of the BDE (for both coarse delay and fine delay)

bit-pattern. As shown in Fig. 5(a), each delay element consists of both fine tune delay block and coarse tune delay block which can be programmed through SPI to match the required chip-length. Along with the tunable delay blocks, a fixed delay value is also present to set the minimum chip-length. In comparison to the published works [21, 22], the presented DE are very simple CMOS implementation without any complex calibration circuits.

The implementation of fine tune or coarse tune delay with schematic is as shown in Fig. 5(b). The basic delay elements (BDE)s are cascaded and connected to a multiplexer to form the delay block. The number of DE determines the programmed delay setting. For the fine tune block, BDE consists of an inverter to provide the fine-grain delay tuning with a resolution of around 0.05 ns and can be tuned to a maximum value of 1.6 ns. For unit delays with chip length of ≈ 20 ns, we expect 0.05 ns resolution to be sufficient. The BDE in the coarse tune block is an inverter with cascaded NMOS and PMOS transistors to provide the required coarse delay with minimal area penalty. The resolution of the coarse tune block is around 0.5 ns and can be tuned to a maximum value of 16 ns. Hence the maximum tunable delay value per delay element is ≈ 21.6 ns. The process dependent variation, mismatch and jitter impacts the delay line with respect to PW and chip-interval of the incoming symbol. As the decision time for the pulse is limited by the pulse extender, the mismatch between the taps must be tuned through calibration.

5.2 Correlator

The continuous-time RAKE receiver is based on OOK and hence correlation refers to matching of bit '1's in the incoming signal. The correlator is based on digital AND

gate which continuously correlates the delay line output and detects the number of pulses (bit '1') present in the received symbol. By avoiding fast clocking registers and parallelism, the correlator is very power-efficient and at the same time performs high speed correlation without any additional overhead.

5.3 Counter

The counter consists of a 2-input multiplexers arranged in a matrix form ($N\text{-bit} \times N\text{-bit}$) as shown in Fig. 1. The correlator output are continuously routed to the counter output as a thermometer-coded digital value, which simplifies the readout circuitry.

5.4 Match thresholder

The counter output appears at the match thresholder input as thermometer-coded digital value and by setting the appropriate value in the match thresholder, a statistical symbol estimation is feasible.

6 Measured results

The CTBV ranging system which uses the continuous-time RAKE receiver was designed and fabricated using TSMC

90 nm low power process technology. The chip microphotograph of the CTBV ranging system which includes the continuous-time RAKE receiver and other necessary blocks for ranging are as shown in Fig. 6. The PCB board including a suitable microcontroller along with all other necessary components for the measurement is as shown in Fig. 7. For the measurement purpose, the output circuitry consists of a continuous-time threshold sampler which is already published in [23].

The CTBV ranging system consists of an integrated symbol generator which uses the Gaussian pulse generator [24] to transmit the required symbol pattern. Gold code such as (1100001100100011100110110110101) is used for transmission which has auto-correlation of 16, maximum non-zero-offset auto-correlation value of 8 and maximum cross-correlation value of 8. The measurement setup consists of two PCBs with antennas, one as a transmitter and another in a receiver mode.

6.1 Delay line calibration

The delay line in the symbol detector needs an initial calibration to compensate for static and dynamic variations effecting the chip-interval or pulse position of the incoming signal. The calibration procedure with respect to pulse position variation and PW variation are automated using python programming.

The pulse width measuring (PWM) block as shown in Fig. 8 aids in tuning the pulse position or chip-interval variation of the incoming signal without any hardware complexity. The tuning or calibration procedure is

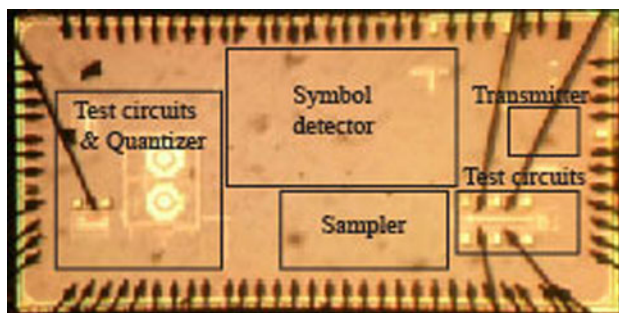


Fig. 6 Chip microphotograph

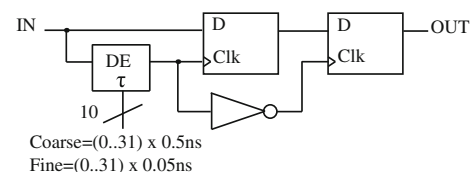


Fig. 8 Schematic of PWM block

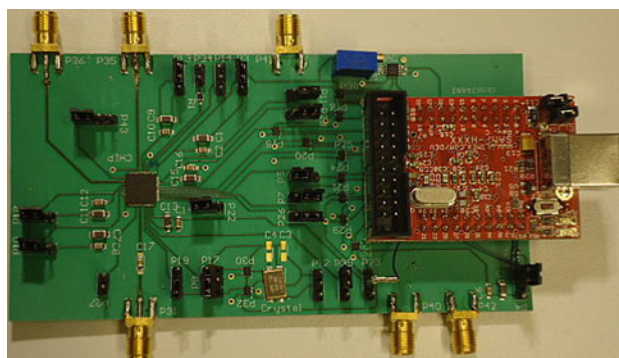


Fig. 7 PCB of the CTBV ranging system

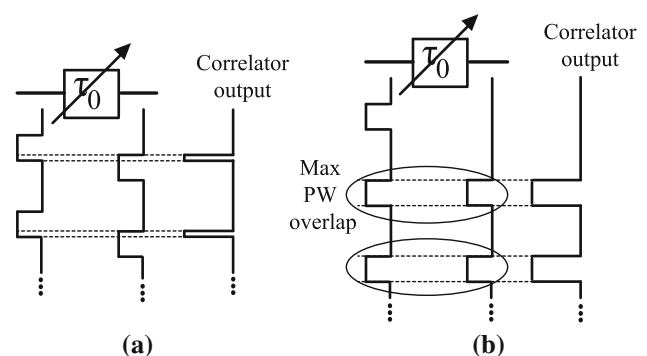


Fig. 9 (a) Without calibration. (b) After calibration

achieved by programming each delay element to the appropriate chip-length by coarse-tune and fine-tune register settings using a SPI interface to the microcontroller. The variation effect on the delay line has an uncorrelated effect on the incoming symbols and its delayed components as shown in Fig. 9(a).

With respect to the first delay element, second delay element is tuned such that they have maximum overlap in PW between their pulses. By enabling just the two correlators, we have two pulses aligned in time in counter output and by setting the match threshold to two, a pulse is detected at the output as shown in Fig. 9(b). The PW overlap between the two correlators is measured using the PWM block. Thus the delay is incremented until the maximum PW overlap is achieved. The calibration technique is measured and plotted as shown in Fig. 10, where it shows that the delays in the x-axis are incremented until the maximum PW overlap (y-axis) is achieved. The plot also shows the variation between the DE to achieve a chip-

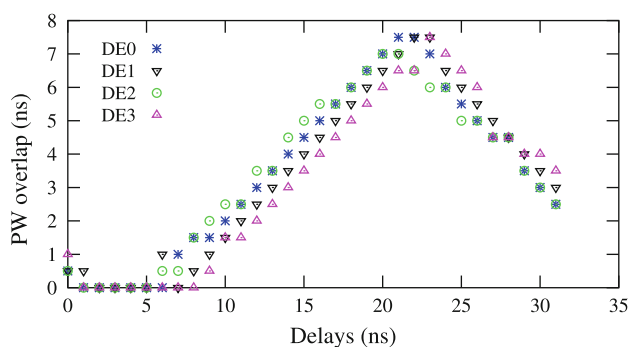


Fig. 10 Sweep of delay values with respect to maximum PW overlap across the DE (DE0, DE1, DE2, and DE3)

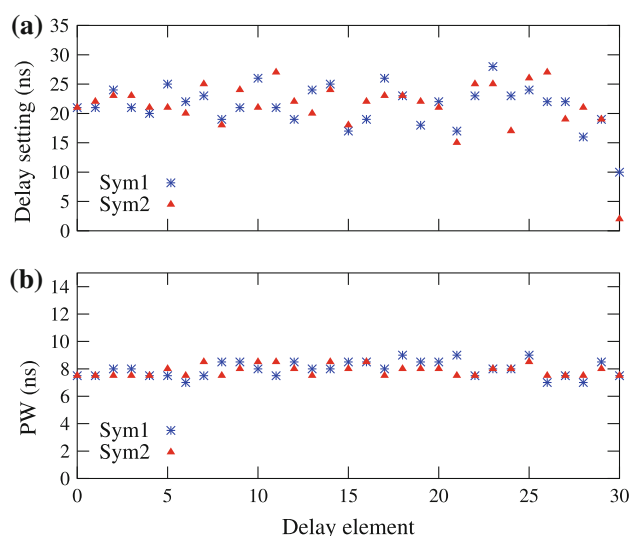


Fig. 11 (a) Calibrated delay values across the DE. (b) Actual PW corresponding to the calibrated delay settings across the DE

interval of 20 ns. For simplicity, DE from 0 to 3 are shown in the plots. In the similar way, all other DE (from 0 to 30) are also tuned with respect to the first delay element so that they have maximum overlap in PW between the DE. Figure 11(a) shows the achieved delays corresponding to maximum PW overlap on the y-axis to set the required chip interval of 20 ns with respect to all the DE after calibration. Figure 11(b) shows the corresponding output PW of the DE after calibration. The PW variation across the DE is between 7 and 10 ns which shows that the DE are calibrated with proper chip-interval and also the correlator output has sufficient PW. The plot also shows that the delays are varying slightly with different symbol patterns (Sym) which could be due to the internal dynamic voltage switching because of the different loading on the power rail caused by the symbol pattern switching through the different signal paths in the long chain of inverters present in the delay line.

The PW variations in the delay line can be significantly reduced by optimal design technique and also the use of pulse width controller blocks (PWC) [20]. The delay line consists of a long chain of multiplexers and inverters, which forms asymmetrical capacitive loading to the signal path. Hence the inverters in the delay element are carefully designed such that their PW variation is constant irrespective of different capacitive loading as shown in Fig. 12. By sweeping the PMOS to NMOS transistor ratio

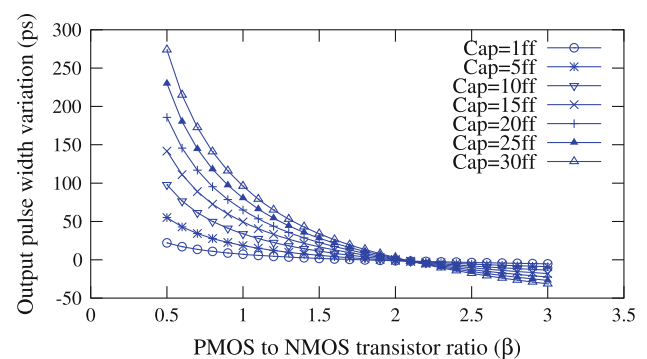


Fig. 12 Optimal design of delay element

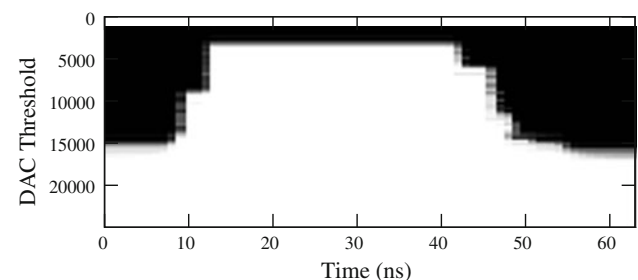


Fig. 13 Image plot of the symbol detected pulse versus DAC range

(β), its corresponding output PW is plotted with different capacitive loading. At a particular value of β , the output PW variation is the same irrespective of the capacitive loading. This particular beta value is used while sizing the BDE. Thus any shrinking or extending in the PW is balanced out by the inverting and non-inverting logic present in the delay line. Further variation in the PW can also be tuned by the PWC in the delay line as illustrated in [20].

6.2 Quantizer threshold settings

After the initial delay line calibration, the quantizer threshold needs to be found. The threshold is swept across the input signal and its corresponding measured image plot using the sampler is as shown in Fig. 13. The plot helps us to estimate a suitable threshold value using the swept threshold technique. It is visible from the plot that the threshold range is between 5,000 and 15,000 DAC values. For further measurements, the DAC value is set to 13,600 which is a good trade-off between range and noise. The plot also shows the range of the switching threshold where stochastic resonance can happen.

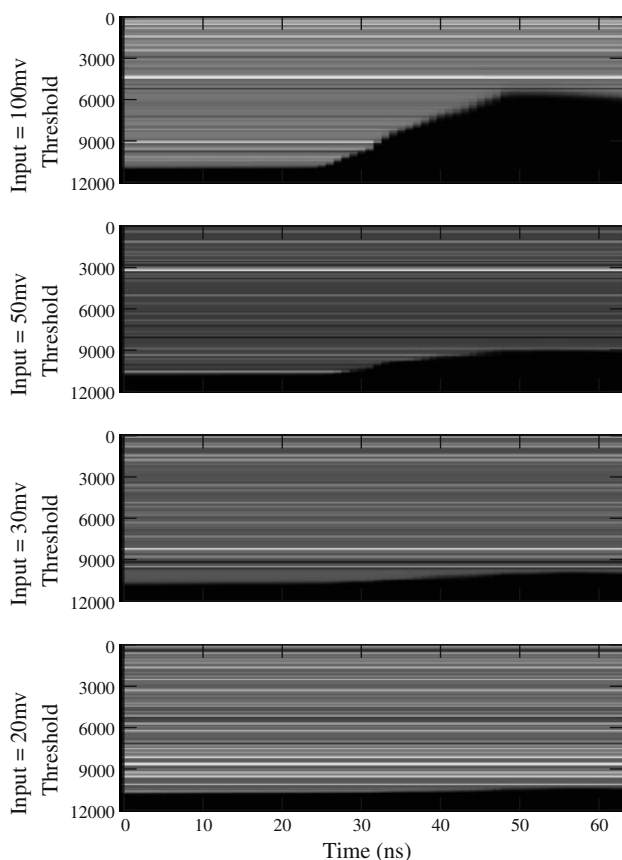


Fig. 14 Symbol detected under different input amplitude

6.3 Symbol detector

With delay line calibration and setting quantizer threshold, the continuous-time symbol detection can be measured. By varying the input signal amplitude, the corresponding symbol detection image plot using the sampler versus threshold range (DAC) is plotted as shown in Fig. 14. The sampler output indicated by white color corresponds to higher counter values whereas the black color corresponds to lower counter values. The plot suggests that symbol detected output is almost buried in the noise when the input amplitude is 20 mV.

The data rate of the continuous-time RAKE receiver is calculated as $1/((T_{\text{chip}} \times N_s + T_{\text{quiet}}) \times n_{\text{repetition}})$. Where T_{chip} is the chip interval and T_{quiet} (≈ 30 ns) is the assumed quiet period after every symbol transmission. $n_{\text{repetition}}$ is the processing gain. N_s corresponds to number of chips. The receiver performance with respect to different chip intervals are as shown in Table 1. The chip interval can be varied to increase data rate but processing gain which is very much needed for the continuous-time system will reduce the data rate. The power dissipation is estimated with respect to 31-chip incoming symbol. With respect to different chip intervals, the power dissipation in symbol detector is minimal whereas the quantizer block is increasing with increased chip interval. The symbol detector is quite power-efficient compared to the quantizer. Symbol detector was custom designed for power-efficient but the quantizer should be optimized in future for low-power.

Table 1 Performance summary

Chip interval (ns)	Data rate (Mbit/s)	Power (mW)	
		SD	Quantizer
18	1.65	0.96	4.29
20	1.5	0.97	4.56
22	1.37	0.97	4.58

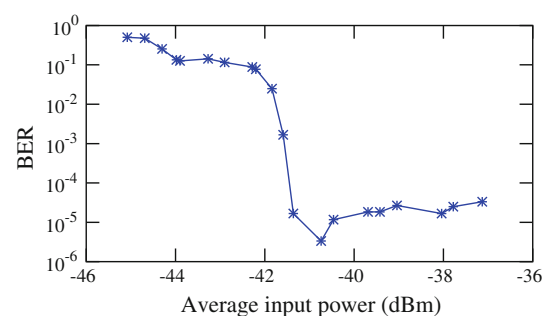


Fig. 15 Measured BER versus average input power

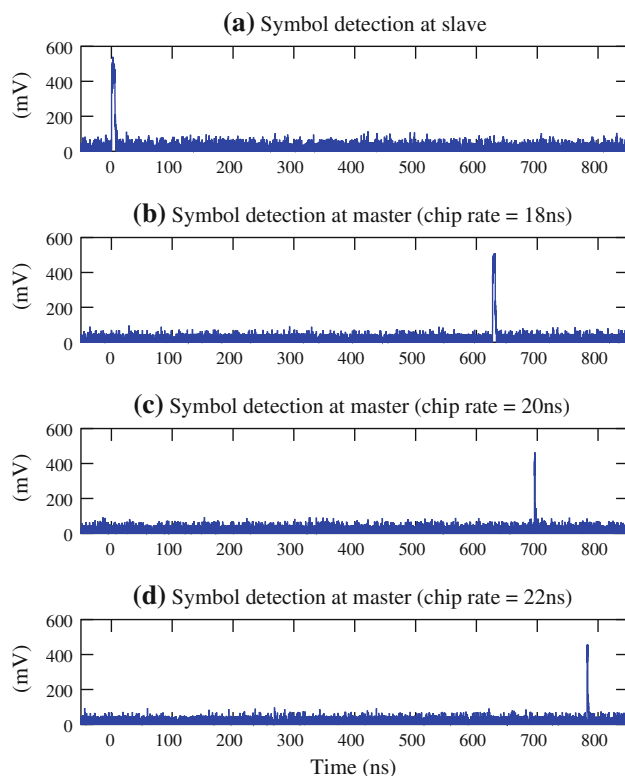


Fig. 16 Continuous-time symbol detection in a master and slave configuration

Table 2 Comparison of continuous-time RAKE receiver

Performance	This work	[11]	[25]
Process (nm)	90	130	130
Area (mm ²)	3.6	6.4	216
Supply voltage (V)	1.2	1.2	1.8–1.2
Architecture	Coherent	Coherent	TDC
Clockless detection	Yes	No	No
Power (mW)	5.29	145.8	46.8

The bit error rate (BER) of the continuous-time RAKE receiver can vary with many factors such as distance between the transmitter and receiver, symbol pattern (code density), processing gain, quantizer threshold and match threshold. The measured BER is plotted against the average input power as shown in Fig. 15. A BER ($\approx 10^{-5}$) is achieved for very short range (<10 cm) which is quite good compared to the long range where noise dominates the signal. The BER rate can be improved further through many ways such as increasing the symbol length, symbol repetition coding and including an LNA in the receiver front-end. The continuous-time RAKE receiver is an initial prototype designed for short-distance ranging system and its coherent symbol detection allows high-precision localization along with communication.

6.4 CTBV ranging system

The continuous-time RAKE receiver along with sampler, ranging block and symbol generator forms the complete CTBV ranging transceiver. The CTBV ranging system in a master-slave configuration is employed for high-precision localization with a reported ranging accuracy of ≈ 1.4 cm [12].

The measurement setup consists of two PCBs one as master transceiver and the other as slave transceiver. The measured continuous-time symbol detection in a master-slave configuration can be depicted as shown in Fig. 16. To initiate the communication and ranging, master is triggered to transmit the symbol and also initializes the ranging and sampler block. Slave will output the detection pulse after the symbol reception as shown in Fig. 16(a). Immediately slave will transmit another symbol and master will output the detection pulse after reception as shown in Fig. 16(b). Now the sampler in the master will have the symbol round trip time (master–slave–master) which is the desired time-of-flight value.

The coherency of the symbol detection is obtained since cross-correlation is running in continuous-time and the detection pulse will be issued immediately when matching occurs. The measured symbol detections for different chip intervals at the master are also plotted in Fig. 16(c, d) which shows its corresponding time-of-arrival. This continuous-time symbol detection feature is used to estimate the precise time-of-flight value without any clock synchronization between the master and the slave. The CTBV ranging system is highly appreciated in short distance ranging such as wireless health-care monitoring systems, which require power-efficient, area-efficient, high-precision ranging and cost-efficient systems.

The receiver performance is summarized and compared to other short-range IR-UWB receivers operating in the 3.1–10.6 GHz UWB spectrum in Table 2. Compared to other works, the presented continuous-time RAKE receiver is very power-efficient and also does coherent clockless symbol detection suitable for high precision ranging.

7 Conclusion

A IR-UWB RAKE receiver utilizing CTBV is presented with measured results. Quantizer threshold which is very critical for the system performance can be accomplished using swept threshold technique. The continuous-time RAKE receiver is clockless and still performs coherent symbol detection without the need for precise timing synchronization. This coherency of the continuous-time RAKE receiver is appreciated in high precision ranging localization. The continuous-time RAKE receiver along with the high-speed sampler, symbol generator and ranging block can be used as a ranging system for high-precision localization ranging without any clock synchronization (≈ 1.4 cm).

Acknowledgment The research is being sponsored by Norwegian Research Council project number 187857/S10 and was carried out at the Nanoelectronics group, Department of Informatics, University of Oslo, Norway.

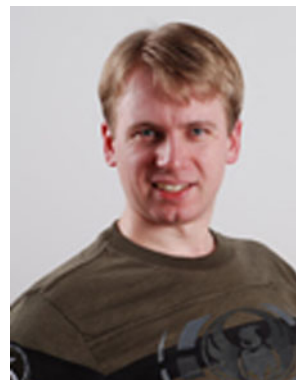
References

1. Ghavami, M., Michael, L., & Kohno, R. (2007). *Ultra wideband signals and systems in communication engineering*. New York: Wiley.
2. Win, M. Z., & Scholtz, R. A. (1998). On the robustness of ultra-wide bandwidth signals in dense multipath environments. *IEEE Communications Letters*, 2(2), 51–53.
3. Win, M. Z., & Scholtz, R. A. (1998). Impulse radio: How it works. *IEEE Communications Letters*, 2(2), 36–38.
4. Zhang, J., Orlik, P., Sahinoglu, Z., Molisch, A., & Kinney, P. (2009). UWB systems for wireless sensor networks. *Proceedings of the IEEE*, 97(2), 313–331.
5. Taylor, J. D. (1995). *Introduction to ultra-wideband radar systems*. Boca Raton: CRC Press.
6. Verhelst, M., & Dehaene, W. (2008). A flexible, ultra-low-energy 35 pJ/pulse digital back-end for a qac IR-UWB receiver. *IEEE Journal of Solid-State Circuits*, 43(7), 1677–1687.
7. Chao, Y.-L., & Scholtz, R. A. (2005). Ultra-wideband transmitted reference systems. *IEEE Transactions on Vehicular Technology*, 54(5), 1556–1569.
8. Lachartre, D., Denis, B., Morche, D., Ouvre, L., Pezzin, M., Piaget, B., Prouvee, J., & Vincent, P. (2009). A 1.1nJ/b 802.15.4a-compliant fully integrated uwb transceiver in 0.13 μ m CMOS. *Solid-state circuits conference-digest of technical papers, 2009. ISSCC 2009. IEEE International* (pp. 312–313, 313a).
9. Verhelst, M., Vereecken, W., Steyaert, M., & Dehaene, W. (2004). Architectures for low power ultra-wideband radio receivers in the 3.1–5 GHz band for data rates <10 mbps. *Proceedings of the international symposium on low power electronics and design*.
10. Van Helleputte, N., Verhelst, M., Dehaene, W., & Gielen, G. (2010). A reconfigurable, 130 nm CMOS 108 pJ/pulse, fully integrated IR-UWB receiver for communication and precise ranging. *IEEE Journal of Solid-State Circuits*, 48(1), 69–83.
11. Zhou, L., Chen, Z., Wang, C.-C., Tzeng, F., Jain, V., & Heydari, P. (2011). A 2-Gb/s 130-nm CMOS RF-correlation-based IR-UWB transceiver front-end. *IEEE Transactions on Microwave Theory and Techniques*.
12. Sudalaiyandi, S., Hjordland, H. A., Vu, T. A., Næss, Øivind, & Lande, T. S. (2012). Continuous-time high-precision IR-UWB ranging-transceiver in 90 nm CMOS. *IEEE Asian solid-state circuits conference*.
13. Li, Y. W. et al. (2005). A continuous-time programmable digital FIR filter. *IEEE custom integrated circuits conference*.
14. Schell, Y., & Tsividis, B. (2008). A continuous-time ADC/DSP/DAC system with no clock and with activity-dependent power dissipation. *IEEE Journal of Solid-State Circuits*.
15. Hjordland, H. A., & Lande, T. S. (2009). CTBV integrated impulse radio design for biomedical applications. *IEEE Bio-medical Circuits and Systems*. Oslo: University of Oslo.
16. Dokania, R., Wang, X., Tallur, S., Dorta-Quinones, C., & Apse, A. (2010). An ultralow-power dual-band UWB impulse radio. *IEEE Transactions on Circuits and Systems II: Express Briefs*. (pp. 541–545).
17. Vu, T. A., Sudalaiyandi, S., Hjordland, H., Naess, O., Lande, T., & Hamran, S. (2012). A variable-gain single-bit ultra-wideband quantizer for baseband receiver front-end. *IEEE Asia Pacific conference on circuits and systems (APCCAS)*, 483–486.
18. Hjordland, H., Wisland, D., Lande, T., Limbodal, C., & Meisal, K. (2007). Thresholded samplers for UWB impulse radar. *IEEE International Symposium on Circuits and Systems*.
19. Bulsara, A. R., & Zador, A. (1996). Threshold detection of wideband signals: A noise-induced maximum in the mutual information. *Physical Review E*, R2185–R2188.
20. Sudalaiyandi, S., Vu, T. A., Hjordland, H. A., Næss, Øivind, & Lande, T. S. (2012). Continuous-time symbol detector for IR-UWB rake receiver in 90 nm CMOS. *IEEE Asian Pacific conference on circuits and systems*.
21. Maymandi-Nejad, M., & Sachdev, M. (2003). A digitally programmable delay element: Design and analysis. *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*. (pp. 871–878).
22. Schell, B., & Tsividis, Y. (2008). A low power tunable delay element suitable for asynchronous delays of burst information. *IEEE Journal of Solid-State Circuits*, 1227–1234.
23. Hjordland, H. A., Wisland, D. T., Lande, T. S., Limbodal, C., & Meisal, K. (2007). Thresholded samplers for UWB impulse radar. *IEEE International Symposium on Circuits & Systems*.
24. Lee, K. K., Dooghabadi, M. Z., Hjordland, H. A., Næss, & Lande, T. S. (2011). A 5.2 pJ/pulse impulse radio pulse generator in 90 nm CMOS. *IEEE International Symposium on Circuits and Systems*.
25. Kang, M. K., & Kim, T. W. (2012). CMOS IR-UWB receiver for ± 9.7 -mm range finding in a multipath environment. *IEEE Transactions on Circuits and Systems II: Express Briefs*. (pp. 538–542).



Shanthi Sudalaiyandi received her Bachelor's degree in Electronics and Communication Engineering from Bangalore University in 2001 and Master's degree in Electrical Engineering from Royal Institute of Technology, Sweden in 2005. From 2005 to 2006, she was a Research Assistant with IMTEK, Germany. In IMTEK, she worked on low power circuit design techniques for WSN. Later she joined Texas Instruments, India as Design Engineer

from 2006 to 2009, wherein she was engaged from standard cell design to chip design activities. Since May 2009, she has been with Nanoelectronics group working towards her Ph.D. degree. Her research interests include digital circuit design and high-precision localization using IR-UWB technology.



Håkon A. Hjordland received the M.Inf. degree in microelectronics from the University of Oslo, Norway in 2006, where he is currently working towards the Ph.D. degree. His research interests include UWB impulse radio, short range radar and wireless data communication.



Tor Sverre Lande IEEE Fellow, is a Professor in the Nanoelectronics at Department of Informatics, University of Oslo. He is a visiting professor at Institute of Biomedical Engineering, Imperial College, London, UK. His primary research is related to microelectronics, both digital and analog. Research fields are Neuromorphic Engineering, analog signal processing, micropower circuit and system design, biomedical circuits and systems and

impulse radio (IR-UWB) for wireless and radar. He is the author or co-author of more than 100 scientific publications with chapters in

two books. He is serving as an associate editor of several scientific journals. He has served as guest editor of special issues like IEEE Transactions on Circuits and Systems, vol II and IEEE Transactions on Circuits and Systems, vol I, special issue on "Biomedical Circuits and Systems". He is/has been a technical committee member of several international conferences and has served as reviewer for a number of international technical journals. He has served as Technical Program Chair for several international conferences (ISCAS2003, NORCHIP, BioCAS, ISCAS2011). He was chair elect (2003–2005) of the IEEE Biomedical Circuits and Systems Technical Committee (BioCAS) and is also a member of other CAS Technical Committee. In 2006 he was appointed Distinguished Lecturer of the IEEE Circuits and Systems Society and elected member of CAS Board of Governors. He is the first Editor-in-Chief of the new IEEE Transactions on Biomedical Circuits and System (2007–2010). He is also the co-founder of the spin-off company, Novelda AS. IEEE Fellow in 2010.

Paper 12 An IR-UWB Transmitter for Ranging Systems

M. Z. Dooghabadi, H. A. Hjortland, O. Nass, K. K. Lee, and T. S. Lande. “An IR-UWB Transmitter for Ranging Systems”. *Circuits and Systems II: Express Briefs, IEEE Transactions on*, Vol. 60, No. 11, pp. 721–725, Nov. 2013. doi:10.1109/TCSII.2013.2281909. © 2013 IEEE. Reprinted with permission.

My Contributions to This Paper

This paper presents a transmitter with continuous-time triggering for the ranging system. It consists mainly of a clock burst generator, a symbol generator, and a pulse generator. I did most of the design for each circuit described in the paper, except for the pulse shaper. My SPI and microcontroller solution was used to control the chip. I also helped review the paper.

An IR-UWB Transmitter for Ranging Systems

Malihe Zarre Dooghabadi, Håkon A. Hjortland, Øivind Næss, Kin Keung Lee, and Tor Sverre Lande, *Fellow, IEEE*

Abstract—This brief presents a continuous-time impulse radio ultrawideband transmitter. The transmitter is a part of a high-precision ranging single-chip transceiver that measures the time-of-flight symbol propagation. The clock burst generator in the transmitter will initiate symbol transmission in continuous time unbounded by any clock signal while maintaining an accurate chip rate during symbol transmission. Using a calibration circuit, the clock period can be programmed precisely to compensate for device mismatch. The transmitter is fabricated in Taiwan Semiconductor Manufacturing Company 90-nm CMOS technology and occupies an area of 0.123 mm². The programmable clock range is from 12.65 to 111 MHz, and the measured rms jitter is 3.26 ps at 50 MHz. The entire transmitter has a power consumption of 1.41 mW at the data rate of 2 Mbit/s.

Index Terms—Calibrated clock, clock burst generator, continuous-time binary value (CTBV), impulse radio ultrawideband (IR-UWB) transmitter, ranging system.

I. INTRODUCTION

ULTRA WIDEBAND (UWB) wireless communication systems may be categorized into two main groups: impulse radio UWB (IR-UWB) systems and multiband orthogonal frequency-division multiplexing systems [1]. The IR-UWB uses the transmission of very short pulses with large bandwidth and low duty cycles, enabling the implementation of a low-power IR-UWB transmitter. Therefore, IR-UWB is well suited for short-range and low-power applications such as radio-frequency identification tags, sensor networks, biomedical electronics, and localization applications.

This brief presents an all-digital IR-UWB transmitter based on the continuous-time binary-value (CTBV) signal processing technique. In the CTBV technique, there is no running clock signal, and timing is based on the inherent process-dependent gate delays [2]. In the CTBV signal processing domain, the signal amplitude is a binary value of either 1 or 0, and it is continuous in time. In the presented transmitter, the on-off keying (OOK) modulation scheme is used because it is less complex and can be realized using simple digital circuits.

The transmitter is intended in a ranging IR-UWB transceiver for measuring the symbol round-trip time [3]. In clocked systems, the time-of-flight (ToF) measurement and the ranging resolution are limited by the clock frequency [4]. In this effort, we achieve improved ranging resolution by doing symbol ToF measurements in continuous time [3]. In the presented

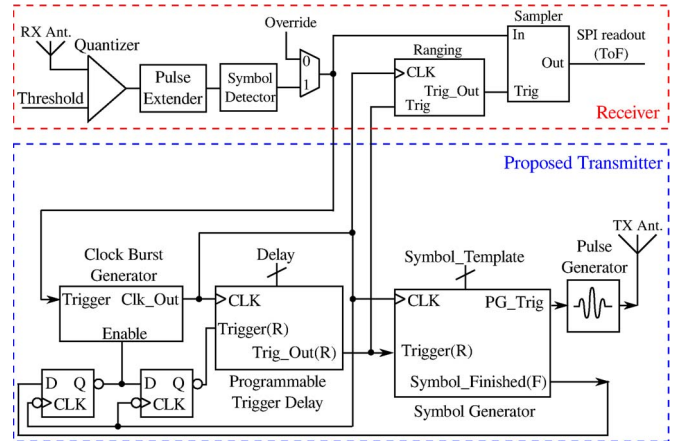


Fig. 1. Ranging transceiver system and the proposed IR-UWB transmitter.

CTBV-based ranging system shown in Fig. 1, we are able to reach a ranging precision of 1.4 cm with moderate design complexity [3].

In the ranging system, the master transceiver sends symbol S1 and initiates the ToF measurement. The slave transceiver detects symbol S1 after some propagation delay. When the symbol detector in the slave transceiver detects S1, then another symbol S2 is returned to the master transceiver. A constant processing time is added since no clock is involved. After some propagation delay, the master transceiver detects S2, and ToF is measured using the high-speed sampler. In the measured symbol round-trip time, all the processing must be constant in time, and the only variable delay is the actual ToF measuring the distance between the transceivers.

In such continuous-time ToF measurements, symbol transmission must be unbounded by any clock signal for proper operation. However, a high-quality periodic signal is required for a chip clock during symbol transmission. In the proposed transmitter, there is a clock burst generator circuit operating in continuous time, generating a preset sequence of clock pulses. The clock burst generator has an immediate start-up triggered by the rising edge of an incoming trigger signal, and it can also be stopped quickly. Therefore, the proposed clock burst generator will always give constant and stable pulse sequences.

One of the important design parameters in ranging systems is a periodic and stable chip rate in order to avoid misalignment and ensure reliable symbol detection. To ensure the correct chip rate and thus reduce the chip rate variability, a calibration circuit and a calibration procedure using an external accurate clock reference (crystal oscillator) are used [5]. The external clock reference is only briefly used in the calibration phase and disabled during the system operation mode. In the calibration mode, the clock burst generator circuit is initially programmed and calibrated precisely.

Manuscript received October 23, 2012; revised April 29, 2013; accepted September 9, 2013. Date of publication September 30, 2013; date of current version November 14, 2013. This work was supported by The Research Council of Norway under Project 187857/S10. This brief was recommended by Associate Editor T. Tong.

The authors are with the Department of Informatics, University of Oslo, 0316 Oslo, Norway (e-mail: malihezdz@ifi.uio.no; bassen@ifi.uio.no).

Color versions of one or more of the figures in this brief are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TCSII.2013.2281909

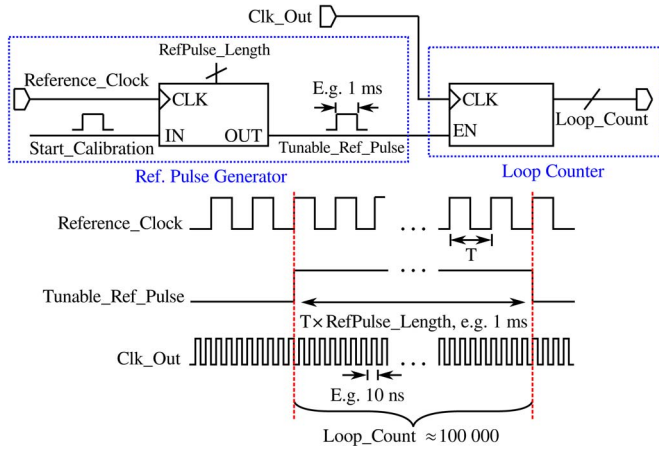


Fig. 2. Clock burst generator calibration circuit.

The rest of this brief is organized as follows. Section II describes the transmitter system operation and architecture. Measurement results are presented in Section III and IV concludes this brief.

II. PROPOSED TRANSMITTER

As shown in Fig. 1, the presented transmitter has four main circuit blocks: a clock burst generator, a programmable trigger delay, a symbol generator, and a pulse generator. The transmitter is triggered by the symbol detector, recovering the incoming UWB signals. On the rising edge of the incoming trigger signal, the clock burst generator starts immediately. Before symbol transmission, a setup time is inserted for stabilization of the oscillation. Since this setup time is dependent on the chip rate, a Serial Peripheral Interface (SPI) programmable register is provided for the number of setup cycles. After the setup time, a pulse sequence is generated according to the programmed symbol pattern. The rising edge of the symbol pattern will then initiate a higher order Gaussian pulse for the antenna [6]. When the symbol transmission is completed, a short pulse is generated on Symbol_Finished(F) to stop the clock burst generator.

Due to the immediate start-up and stop time of the clock burst generator, the transmitter always has a constant processing time, which is a key function of the continuous-time ToF measurement system. There is simply no clock to wait for in continuous-time systems, so the start-up time will always be constant. However, other noise sources like jitter noise are still present.

A. Transmitter Circuit Description

1) *Clock Burst Generator*: The clock burst generator generates a finite sequence of clock cycles. The module consists of two main circuits: a calibration circuit and a clock generator loop. The calibration circuit and the clock burst generator are shown in Figs. 2 and 3, respectively.

The clock burst generator has two operation modes: manual and trigger. In the trigger mode, the clock generator loop is enabled for clock signal generation and in the manual mode, the clock period can be calibrated and programmed precisely using the calibration circuit. First, the manual mode should

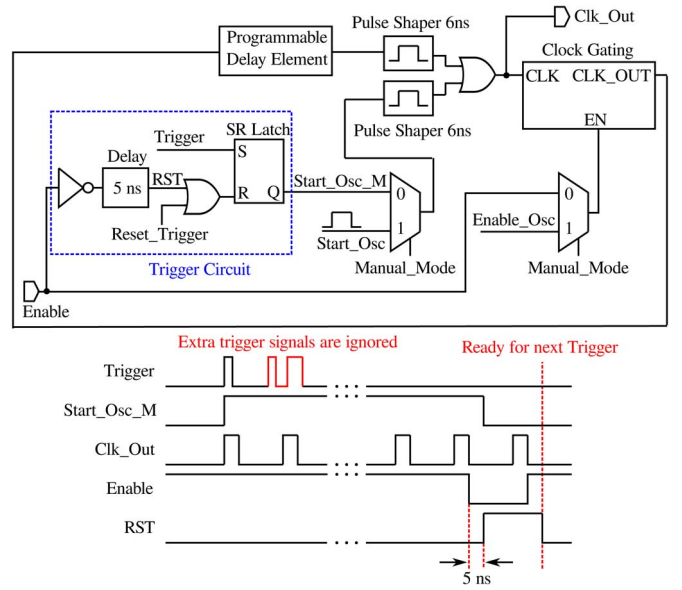


Fig. 3. Clock burst generator loop.

be enabled for calibration of the clock period to the desired value, and after that, the trigger mode is enabled and the clock burst generator is ready to operate. In the following, the manual and trigger modes of the clock burst generator are explained in detail.

a) *Manual mode*: In the manual mode, by setting Enable_Osc to high, the ring oscillator is enabled, and oscillation starts by sending a pulse on Start_Osc. The oscillation should be kept running for a short time until the loop signal (Clk_Out) is stabilized. Then, calibration can be started by sending a pulse on Start_Calibration, and after that, the tunable reference pulse is generated. As long as the reference pulse is high, the loop counter counts the oscillation cycles of the clock generator loop signal Clk_Out. The period of the loop signal is the reference pulsewidth divided by the loop counter value, e.g., for a tunable reference pulse of 1 ms and a loop count of 100 000, the period of the loop signal is 10 ns. Thus, by tuning the programmable delay element inside the loop, a clock signal with a given period can be generated. The timing diagram in Fig. 2 shows the measurement procedure.

b) *Trigger mode*: The timing diagram in Fig. 3 demonstrates the circuit operation in trigger mode. In the trigger mode, Enable must be set high for two purposes: first for enabling the ring oscillator and second for enabling the triggering of the oscillation. The inverted and delayed Enable makes the reset signal (RST) for the SR latch (R prioritized before S). The SR latch can also be reset using the manual reset signal (Reset_Trigger).

When Enable is set high, the first incoming input pulse on Trigger will change the output of the SR latch from low to high, and then, the generated rising edge on Start_Osc_M goes to the pulse shaper and starts the oscillation. Any additional pulses on Trigger will not have any effect since the SR latch is already set high. Oscillation is going on as long as Enable is high. After Enable is set low, oscillation will stop after one clock cycle. When Enable is set low, 5 ns later, RST and R will be set high and reset the SR latch. After Enable is set high again,

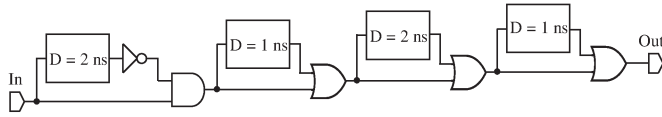


Fig. 4. Pulse shaper circuit.

the ring oscillator is ready to oscillate with the first incoming pulse on Trigger. As shown in Fig. 3, the proposed clock burst generator starts immediately with the rising edge of the trigger signal, so this circuit can also be used as an immediate clock synchronizer.

c) *Clock burst generator design consideration:* The clock generator loop is a noninverting ring oscillator and has four parts: a pulse shaper, a clock gating circuit, a trigger circuit, and a programmable delay element.

In a ring oscillator, tuning the delay value of the loop delay element changes both the rising and falling edge propagation delays. In a conventional ring oscillator, both the rising and the falling edge propagation delays determine the loop signal period, so if the delay value of the loop delay elements is increased or decreased with 50 ps, the loop signal period will change about 100 ps. However, in a noninverting ring oscillator, only the rising edge propagation delay will affect the loop signal period, thus giving us a finer tuning step size than an inverting ring oscillator. This is simply achieved by inserting a one-shot circuit in the loop.

Fig. 4 shows the pulse shaper circuit. The pulse shaper maintains oscillation inside the loop, and it is a one-shot circuit accepting any input signal with pulsewidth exceeding 2 ns and giving an output signal with a fixed pulsewidth of 6 ns. Since the clock period of the loop signal is determined by the rising edge propagation delay, the pulsewidth variation is not a critical issue. In the presented design, the pulsewidth is set to 6 ns, which is about half of the minimum of the loop signal period.

The clock gating circuit is used for enabling and disabling the ring oscillator, facilitating a programmable sequence of pulses. The trigger circuit injects a 6-ns pulse into the ring oscillator at the first rising edge of the input trigger signal, thus starting the ring oscillator. In the trigger circuit, when the Enable signal is set low, 5 ns later, the RST signal is generated to make sure that there is at least 5-ns low time between the last clock pulse of one clock burst and the first clock pulse of the next clock burst. Without a pause time of 5 ns, which is about half of the minimum of the loop signal period, there can be a risk of getting two clock pulses right after, which may create unexpected behavior in the rest of the system.

The programmable delay element is a cascade of three different programmable subelements, called fine, medium, and coarse delay lines. The fine delay element has 65 delay settings with tuning steps of 10 ps. The medium and coarse delay lines have 64 delay settings with step sizes of 50 ps and 1 ns, receptively. By tuning the programmable delay element, the clock period can be set to the desired value.

2) *Programmable Trigger Delay:* The programmable trigger delay is shown in Fig. 5. This circuit is a coarse delay element with delay steps of one clock period. The required delay value is set by programming the Delay register. The circuit is triggered by the incoming trigger signal, which sets

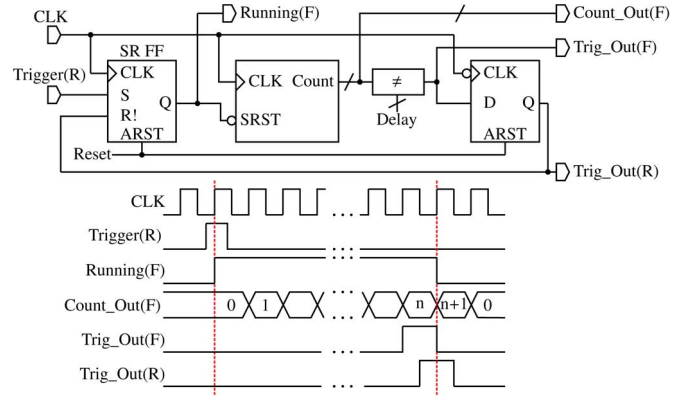


Fig. 5. Programmable trigger delay.

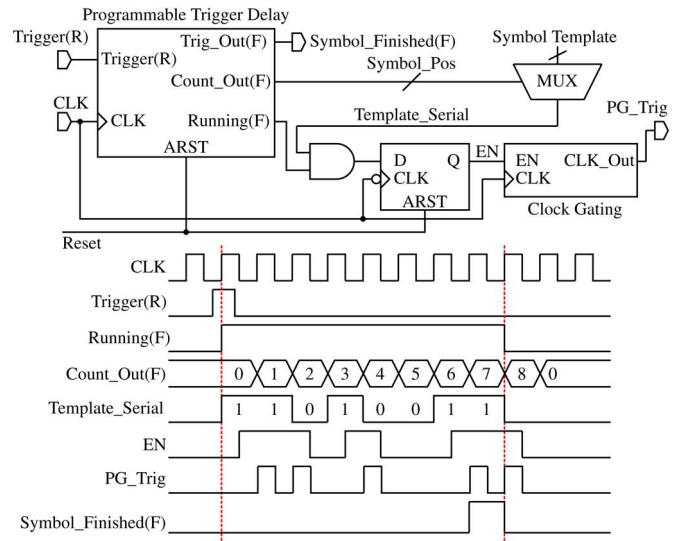


Fig. 6. Symbol generator.

the Running(F) signal high, and starts the counter. When the counter increments up to the preset delay value stored in the Delay register, the Running(F) signal goes low, and the trigger output signals will be generated on Trig_Out(F) and Trig_Out(R). The timing diagram in Fig. 5 illustrates the circuit functionality; as shown, if the Delay register is set to n , then the Running(F) signal stays high for $n + 1$ and the counter counts to $n + 1$, so the actual delay value is $n + 1$.

3) *Symbol Generator:* Fig. 6 shows the symbol generator circuit. This circuit generates a sequence of trigger signals according to the desired symbol pattern. The OOK modulation scheme is used in the presented transmitter, so for each of the bits set in the symbol pattern, a trigger pulse will be generated for the pulse generator.

The symbol pattern can be programmed using the Symbol Template register, and the symbol length can be set using the Delay register inside the programmable trigger delay circuit. When the symbol generator is triggered, the symbol is sent serially on the Template_Serial and PG_Trigger signals according to the programmed symbol pattern. When the last bit of the symbol is sent on Template_Serial, a one-clock-cycle pulse will be generated on Symbol_Finished(F). The timing diagram is shown in Fig. 6.

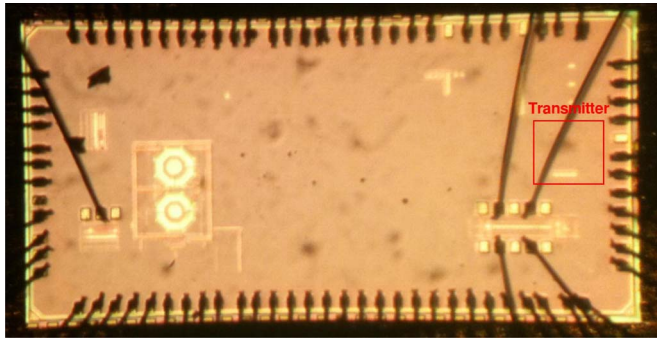


Fig. 7. Chip die photograph.

4) *Pulse Generator*: The pulse generator that is presented in [6] is used in the transmitter. The pulse generator generates high-order Gaussian pulses. Several loose triangular pulses are used to construct a first-order Gaussian envelope. The bandwidth is determined by the envelope. The generated Gaussian pulse has a center frequency of about 4 GHz and a bandwidth of about 2.3 GHz, which fits the Federal Communications Commission (FCC) mask [6].

III. MEASUREMENT RESULTS

The proposed transmitter was implemented in the Taiwan Semiconductor Manufacturing Company 90-nm CMOS technology. The die photograph of the ranging transceiver is shown in Fig. 7. The prototype chip has an SPI that is used for setting the registers and generating the controlling signals. In the measurement setup, a microcontroller programmed through the Universal Serial Bus interface is used for signal generation, SPI programming, data collection, and readout.

A. Measured Performance of the Clock Burst Generator

An external reference clock signal with a frequency of 3 MHz is used for calibration and period measurement, and the tunable reference pulse is set to 1 ms. As mentioned in Section II-A1, the pulse shaper may be triggered by any signal with a pulsewidth of more than 2 ns. The pulse shaper also needs a signal with a minimum low time of 2 ns. Thus, the loop delay should always be about 10 ns or more for the circuit to work properly. By tuning the fine, medium, and coarse delay lines in the programmable delay element, the chip rate can be calibrated from 12.65 to 111 MHz.

Not only the clock period of the clock signal is determined by the programmable delay element but also the other digital gates inside the oscillation loop contribute to the total delay of the loop signal. The accuracy of the calibration is determined by the amount of averaging, i.e., the width of the reference pulse [5]. By calibrating the loop delay elements, the process variations and device mismatch can be compensated, but there may still be some clock frequency drifting due to the dynamic noise, e.g., jitter and substrate coupling. The measured fine tuning step is about 10 ps, so we have a limited tuning resolution. For example, with a clock frequency of 100 MHz, we may have an error up to 500 ppm. The clock jitter is measured using the histogram analysis on the Agilent 54855A Infiniium

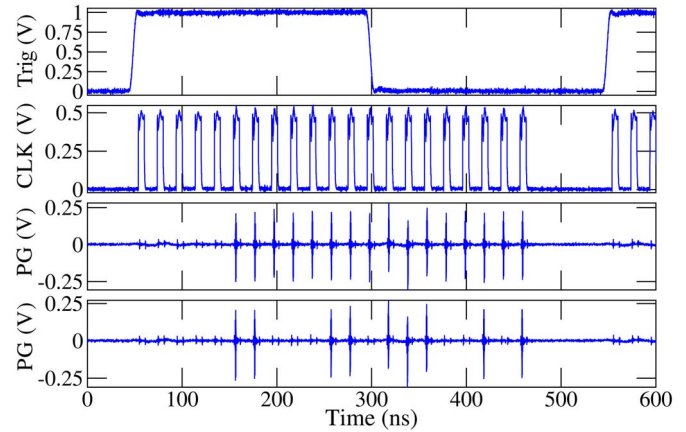


Fig. 8. First graph shows the external trigger signal. The second, third, and fourth graphs show the measured waveforms at the outputs of the clock burst generator and pulse generator. The symbol patterns in the third and fourth graphs are 1111111111111111 and 1100011011100101, respectively.

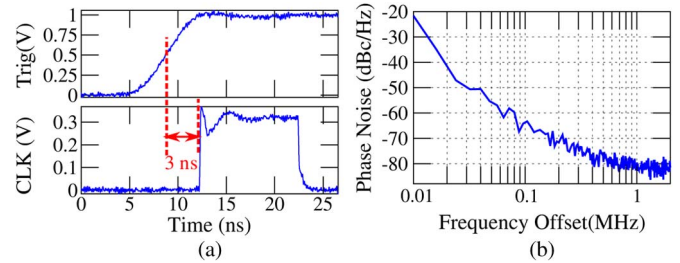


Fig. 9. Measured clock signal and its phase noise at 50 MHz.

oscilloscope. The measured peak-to-peak and rms jitters are 26 and 3.26 ps at a clock frequency of 50 MHz and also 20 and 3 ps at a clock frequency of 100 MHz, respectively.

B. Measured Performance of the Transmitter

The clock period is calibrated and set to 50 MHz to measure the transmitter performance. For measurement purposes, an external square-wave signal is used to trigger the transmitter. The frequency of the square-wave signal is set to 2 MHz in order to have a data rate of 2 Mbit/s. The Delay register in the programmable trigger delay is set to three; therefore, the setup time lasts four clock cycles, and then one clock cycle later, the symbol generator starts triggering the pulse generator. As shown in Fig. 8, after five clock cycles, the pulse generator generates pulses according to the programmed symbol pattern. The symbol length is set to 16. Fig. 8 shows the input trigger signal and the measured waveforms at the outputs of the clock burst generator as well as the pulse generator. The third and fourth graphs in Fig. 8 present two different programmed symbol patterns.

The operation range of a ranging system depends on the transceiver sensitivity, which can be improved by increasing the transmit power, unfortunately severely limited by FCC regulation. Alternative sensitivity improvements of pulsed radio systems would involve some sort of processing gain like symbol repetition coding or extended symbol length trading in transfer rate.

TABLE I
PERFORMANCE SUMMARY AND COMPARISON

Comparison of Transmitters							
	Process	Supply	Modulation	Chip rate	Center Freq.	Power (mW)	Area (mm ²)
This work	90 nm	1.2 V	OOK	12.65-111(MHz)	4.05 GHz	1.41@2Mb/s	0.123
[7]	90 nm	0.9-1.1 V	OOK, S-OOK	—	2.9–3.8 GHz	0.258@2Mb/s	0.6 (incl. Pads)
[8]	180 nm	1.2-1.8 V	OOK, BPSK	—	4 GHz	0.92@1Mb/s	—
[9]	90 nm	1 V	PPM, BPSK	—	3.5 GHz	4.36@15.6Mb/s	0.07
Comparison of Clock Generators							
	Process, Supply	Architecture	Lock-Time, Start-Up	Operation Range	Jitter pk-pk (ps)	Power (mW)	Area (mm ²)
This work	90 nm, 1.2 V	open loop	—, 5 cycles	12.65-111 MHz	20 @100 MHz	3.24 @100 MHz	0.048
[10]	180 nm, 1.8 V	open loop	4 cycles, —	0.1-1 GHz	12 @100 MHz	16 @100 MHz	0.21
[11]	180 nm, 1.8 V	open loop	2 cycles, —	0.8-1.6 GHz	7 @1 GHz	17 @1.6 GHz	0.02

As can be seen in Figs. 8 and 9(a), on the rising edge of the trigger signal, the clock burst generator starts to generate the periodic signal without any varying setup time. Fig. 8 shows that, after the symbol is transmitted completely, the clock burst generator turns off, draining minimal power. In a real transmission, the intersymbol spacing would be increased, reducing intersymbol interference (ISI).

An on-chip inverter-based buffer is connected to the output of the clock burst generator for measurement purposes. Due to the gate delay of this output buffer and digital gates, there is a delay difference of about 3 ns between the measured trigger and the clock signal at the output of the buffer, as presented in Fig. 9(a). The clock signal amplitude is less than VDD because of the 50-Ω loading of the oscilloscope on the output buffer. Fig. 9(b) shows the measured phase noise at the clock frequency of 50 MHz. The phase noise is only relevant for the duration of one symbol (relative phase noise). For moderate symbol lengths, a continuous-time correlator combined with pulse extenders in the receiver is compensating for the added phase noise as proved by the measured performance of the ranging transceiver in [3].

In order to measure the power consumption of the clock burst generator, the signal transmission is turned off, and the clock burst generator is programmed to generate a clock signal continuously. At clock frequencies of 50 and 100 MHz, the clock burst generator has average power consumptions of 1.68 and 3.24 mW, respectively. The presented clock burst generator has less power consumption compared to phase locked loop (PLL)- or delay locked loop (DLL)-based clock generator circuits. As explained in Section II-A1, the programmable delay element inside the clock burst generator is a cascade of three different kinds of delay elements. In the prototype clock burst generator, the unused delay modules inside the programmable delay element are not disabled and significantly contribute to the power consumption. In future design, this can be changed for improved power performance. To measure the power consumption of the transmitter, the symbol pattern is set to 10101010101010, and the data rate and chip rate are set to 2 Mbit/s and 50 MHz, respectively. The measured average current of the entire transmitter including the clock burst generator is about 1.18 mA, so the transmitter has a power consumption of 1.41 mW.

A comparison with similar efforts found in the literature is shown in Table I, indicating improved power performance and reduced chip area.

IV. CONCLUSION

A novel calibrated transmitter for continuous-time symbol generation in IR-UWB communication systems has been presented. The proposed transmitter is a simple delay-based circuit. There is no continuous running clock, and a new calibrated clock burst generator is used to generate a preset periodic pulse sequence for symbol generation. The clock burst generator is initiated in continuous time for proper operation of the high-precision ranging transceiver. By elaborating calibration, a periodic clock sequence is generated for symbol generation. The transmitter has been successfully demonstrated in ranging communication systems [3], giving close to 1-cm accuracy. With good power efficiency and minimal standby power, the reported transmitter is suitable for battery-operated wireless localization systems.

REFERENCES

- [1] J. R. Fernandes and D. Wentzloff, "Recent advances in IR-UWB transceivers: An overview," in *Proc. IEEE ISCAS*, 2010, pp. 3284–3287.
- [2] H. A. Hjortland and T. S. Lande, "CTBV integrated impulse radio design for biomedical applications," *IEEE Trans. Biomedical Circuits Syst.*, vol. 3, no. 2, pp. 79–88, Apr. 2009.
- [3] S. Sudalaiyandi, H. A. Hjortland, T.-A. Vu, O. Naess, and T. S. Lande, "Continuous-time high-precision IR-UWB ranging-transceiver in 90 nm CMOS," in *Proc. IEEE Asian Solid-State Circuits Conf.*, Nov. 2012, pp. 349–352.
- [4] S. Gezici, Z. Tian, G. B. Giannakis, H. Kobayashi, A. F. Molisch, H. V. Poor, and Z. Sahinoglu, "Localization via ultra-wideband radios: A look at positioning aspects for future sensor networks," *IEEE Signal Process. Mag.*, vol. 22, no. 4, pp. 70–84, Jul. 2005.
- [5] M. Z. Dooghabadi, H. A. Hjortland, and T. S. Lande, "High precision calibrated digital delay element," *Electron. Lett.*, vol. 47, no. 9, pp. 564–565, Apr. 2011.
- [6] K. K. Lee, M. Z. Dooghabadi, H. A. Hjortland, O. Nss, and T. S. Lande, "A 5.2 pJ/pulse impulse radio pulse generator in 90 nm CMOS," in *Proc. IEEE ISCAS*, 2011, pp. 1299–1302.
- [7] M. Crepaldi, C. Li, J. R. Fernandes, and P. R. Kinget, "An ultra-wideband impulse-radio transceiver chipset using synchronized-OOK modulation," *IEEE J. Solid-State Circuits*, vol. 46, no. 10, pp. 2284–2299, Oct. 2011.
- [8] Y. J. Zheng, S.-X. Diao, C.-W. Ang, Y. Gao, F.-C. Choong, Z. Chen, X. Liu, Y.-S. Wang, X.-J. Yuan, and C. H. Heng, "A 0.92/5.3nJ/b UWB impulse radio SoC for communication and localization," in *Proc. ISSCC Dig. Tech. Papers*, Feb. 2010, pp. 230–231.
- [9] P. P. Mercier, D. C. Daly, and A. P. Chandrakasan, "An energy-efficient all-digital UWB transmitter employing dual capacitively-coupled pulse-shaping drivers," *IEEE J. Solid-State Circuits*, vol. 44, no. 6, pp. 1679–1688, Jun. 2009.
- [10] M. J. Kim and L. S. Kim, "A 100 MHz-to-1 GHz fast-lock synchronous clock generator with DCC for mobile applications," *IEEE Trans. Circuits Syst. II*, vol. 58, no. 8, pp. 477–481, Aug. 2011.
- [11] M.-Y. Kim, D. Shin, H. Chae, and C. Kim, "A low-jitter open-loop all-digital clock generator with two-cycle lock-time," *IEEE Trans. VLSI Syst.*, vol. 17, no. 10, pp. 1461–1469, Oct. 2009.

A.3 Beamforming

Paper 13 Close range impulse radio beamformers

O. Dahl, H. A. Hjortland, T. S. Lande, and D. T. Wisland. “Close range impulse radio beamformers”. In: *Ultra-Wideband, 2009. ICUWB 2009. IEEE International Conference on*, pp. 205–209, Vancouver, BC, Sep. 2009. doi:10.1109/ICUWB.2009.5288755. © 2009 IEEE. Reprinted with permission.

My Contributions to This Paper

As co-supervisor I worked with the MS student to bring forward a programmable delay element suitable for beamforming of impulse radio solutions. The delay element with back-gate tuning was my idea, and my SPI solution and measurement setup was used.

Close Range Impulse Radio Beamformers

Øyvind Dahl, Håkon A. Hjortland, Tor Sverre “Bassen” Lande and Dag T. Wisland

Dept. of Informatics, University of Oslo, Norway

Email: oyvind.dahl@fys.uio.no, haakoh@ifi.uio.no, bassen@ifi.uio.no, dagwis@ifi.uio.no

Abstract—Beamforming electromagnetic waves, especially at close range, is challenging. Extremely accurate control of phase/timing is required. In this work we present a digital programmable delay-line suited for transmission beamforming using multiple senders. Back-gate tuning of inverter delays are measured to give relative temporal tuning in the range of 1.1 ps.

I. INTRODUCTION

The emergence of Ultra Wideband (UWB) technology in the marketplace is slower than anticipated when the UWB band was released by FCC several years ago. A major obstacle has been to implement solutions in standard CMOS technology. Only a few single-chip CMOS transceiver solutions are currently available either for multi-band or impulse radio [1], [2] and lately several journalists have announced UWB to be dead [3]. In the end it all boils down to cost so single-chip, low-cost CMOS solutions may have significant impact on the UWB market.

At the Nanoelectronics group at Dept. of Informatics, University of Oslo we are pursuing unconventional design paradigms enabling impulse radio solutions in standard digital CMOS operating at microwave frequencies. The Continuous-Time Binary-Value (CTBV) design paradigm has been explored to make a fully functional, single chip impulse radar [4]. A more elaborate treatment of CTBV circuit design for CMOS impulse radio may be found in [5]. An important feature of CTBV circuits is the exploration of inherent gate delay instead of rigorous and power-demanding clocked schemes. One good CTBV example is the high-speed sampler of the impulse radar measuring backscattered energy for every 5th millimeters in depth. This is equivalent to a sampling rate of more than 30 GHz. The short-range CTBV radar is giving remarkably good results using coherent integration. However, unless focused antennas are used, the radiated pulses are spread widely giving backscattered energy from all over the place. In some applications like level measurements the lack of directionality is desirable, but for more detailed information, a controllable electromagnetic (EM) beam would facilitate something like an electromagnetic camera [6]. A focused horn antenna used for scanning would require some mechanical rig but techniques known as beamforming in two dimensions would enable a scannable EM beam. Most EM beamformers are found in radar applications and are named phased arrays. By adjusting the relative phase differences of signals transmitted by several antennas, constructive and destructive interference will set up a radiation pattern a process known as beamforming. A simple time-domain solution is to introduce suitable delays in the sig-

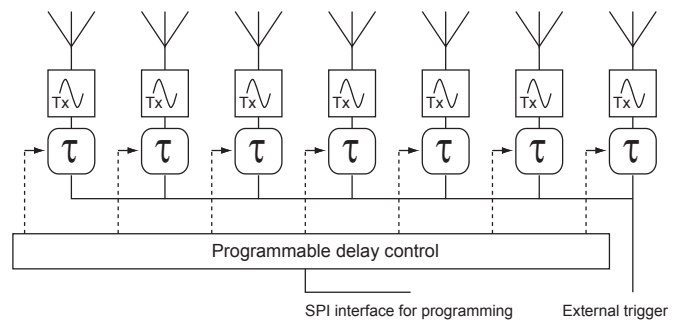


Fig. 1. The transmitter beamforming is achieved by delayed trigger of pulse transmitters.

nal paths before transmission. Although most beamformers are adjusting phases of some sinusoidal signal, we are exploring accurate adjustment of the “firing” sequence of Gaussian pulse transmitters.

Another major challenge doing impulse radio solutions in standard CMOS is the low power supply, approaching 1 V in nanometer technology. Even with rail-to-rail operation, it is hard to radiate the permitted energy level set by the US FCC part 15 limit of -41.3 dBm/MHz. Again adding antenna gain would improve directional power, but beamforming would allow for simple and small antennas. With very power-efficient and low-device-count Gaussian transmitters feasible in CMOS technology [2], adding several on-chip transmitters is simple. By beamforming, significant increase of directional, radiated power is feasible without increasing on-chip power supply. Multiple-antenna arrangements are often named MIMO (multiple-input multiple-output) in communication systems.

A. Organization of the paper

In the start of section II, an overview of the beamforming system is given. In section II-A, the essential part of the system, the picosecond-range programmable delay element, is described. In section II-B, the design of our test-chip, with a fully functional beamformer utilizing the programmable delay, is presented. In section II-C, Monte Carlo simulations and measurements of the statistical behavior of the delay elements is presented. The measurement setup is also described in some detail, and on-chip interference is investigated. In section III, we draw an optimistic conclusion about the presented work.

II. IMPULSE RADIO BEAMFORMER

As shown in Fig. 1 the idea of impulse radio beamforming is simple. A number of identical Gaussian transmitters are

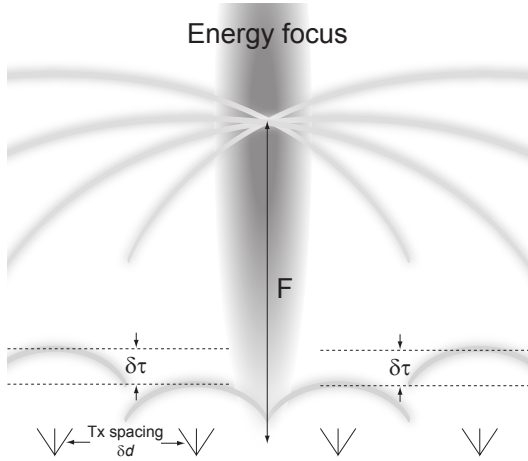


Fig. 2. The radiation from several transmitters will interfere and build up energy when coordinated.

individually controllable with a trigger pulse. The power efficient transmitters employing only a handful of transistors are provided by Novelda AS [2] as part of a technology exchange agreement with Dept. of Informatics, University of Oslo. On the rising edge of the trigger, a Gaussian dual-slope pulse is transmitted on the antenna.

In order to generate a suitable sequence of transmitted pulses, a delay element is inserted between a global trigger pulse and each of the transmitters. In order to focus an EM beam the antennas must be arranged with a suitable spacing. For a given focal distance F and an antenna spacing of δd the required temporal correction is given by

$$\delta\tau = \frac{\sqrt{F^2 + \delta d^2} - F}{c},$$

when we assume signal propagation close to the speed of light $c = 3 \cdot 10^8$ m/s and aligned antennas. For an energy focus at 1 m with antennas spaced 0.1 m apart, the $\delta t \approx 16.6$ ps, while an EM focus at 0.9 m would require $\delta t \approx 18.5$ ps. In this antenna configuration an adjustment of 0.1 m requires a timing adjustment of approximately 1.9 ps. These fine time adjustments are hard to make, especially with the inherent transistor mismatch found in advanced technology. The spatial size of the focused energy is determined by frequency content of the transmitted pulses. In general, higher center frequency will improve focus.

A. Programmable Delay Element

Based on the proposed time-domain impulse radio beamforming, an accurate and programmable delay element tunable in the picoseconds range is required. Even for nanometer technology this is really challenging. The nominal inverter delay is estimated to be approximately 12 ps. An important observation is by using a global edge-triggered topology, we are left with tuning the *relative* time-differences. Adding an overall delay after triggering is acceptable, but the relative delays between the delay-lines must be adjusted accurately in the picoseconds

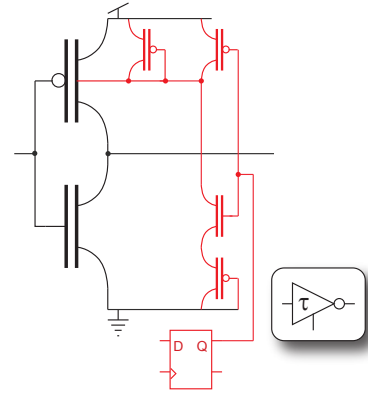


Fig. 3. Binary back-gate inverter tuning using a weak pulldown over a PMOS diode-connected transistor.

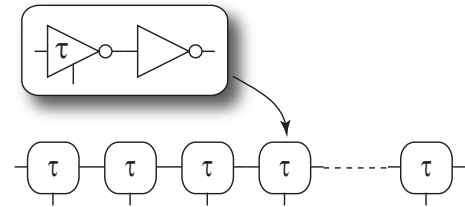


Fig. 4. A programmable delay is constructed by cascading inverters and tunable inverters.

range. So slightly manipulating the individual inverter delays of a longer delay-line is sufficient. Now, cascading several inverters adding latency is acceptable as long as all delays are balanced to approximately the same accumulated time. Minor changes of inverter delays may be done by design with transistor sizing, but a programmable delay is harder to come about. In this approach we are exploring back-gate tuning or body-biasing. Back-gate efficiency is reduced in nanometer technology [7], which in this application is traded in as an advantage enabling fine-tuning of inverter delays.

The circuit topology explored in this work is a simple binary tuning of a PMOS back-gate as shown in Fig. 3. An inverter with weak pull-down is used to switch the body-bias between two voltages. When the inverter output is high, the diode-connected PMOS transistor will drive the well potential to V_{dd} for normal inverter operation. When the inverter is pulling down through the series PMOS load, the well-potential or back-gate is pulled downwards approximately one diode offset, imposing a slight decrease of the front-gate threshold voltage. The effect is a slightly faster PMOS. With a triple-well process, a similar procedure might be used for the NMOS as well, but we are working with edge-triggered transmitters and only need to adjust the positive-going transition. The inverter state is set digitally by storing a bit in a register.

As shown in Fig. 4 the tunable inverter is followed by a normal inverter for signal reconstruction and state inversion. Then a suitable number of delay stages are cascaded to establish required tuning range.

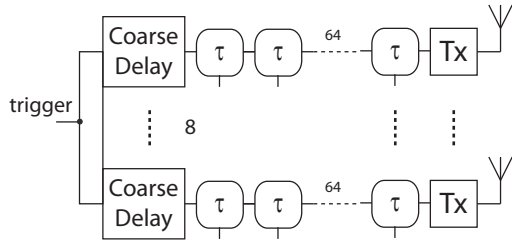


Fig. 5. The evaluation chip consists of eight dual-slope Gaussian pulse transmitters preceded by programmable delay-lines.

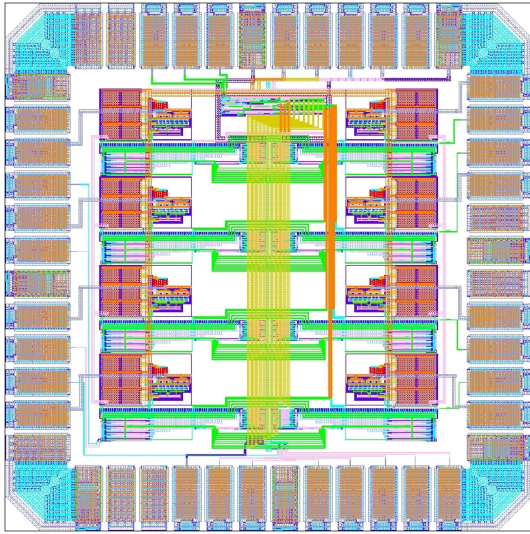


Fig. 6. A chip was designed in STMicroelectronics 90 nm technology with eight transmitters each equipped with programmable delay-lines.

B. Impulse Radio Beamformer Chip

In order to evaluate the quality of delay-line programming a test-chip was made in ST 90 nm CMOS process. Eight identical, edge-triggered impulse radio transmitters for dual-slope Gaussian-shaped pulse transmission were used.

As shown in Fig. 5, the firing of each transmitter is controlled by a programmable delay-line consisting of one programmable coarse-tune delay element and 64 programmable fine-tune delay elements. The coarse delay element is made of a chain of 64 dual inverter delays. Each of the delay-outputs are tapped to a multiplexer for control of the delay-value. The multiplexer enables us to choose the number of dual inverter delays to cascade. The unit delay of the dual inverters are approximately 35 ps, which gives us a programming range of 0 to $63 \times 35 \text{ ps} \approx 2.2 \text{ ns}$. The firing sequence is triggered by an external trigger for all transmitters and a simple SPI scheme is used for programming the different delay-elements integrated on the chip. Fig. 6 shows the chip topology, where the eight transmitters are clearly visible along the pad-frame.

C. Simulated and Measured Performance

As transistor mismatch in fine-pitch processes is significant, a statistical treatment of chip performance is required. As a

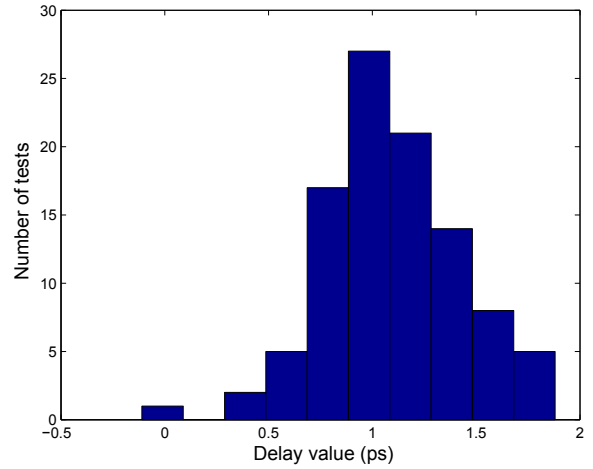


Fig. 7. Simulated probability distribution of the value of one delay-element.

first approximation a delay-element was simulated in Cadence using foundry parameters for Monte Carlo modeling.

The probability distribution in Fig. 7 gives an expected tunable delay of approximately 1 ps. More accurately, we have a mean value of $\mu = 1.1 \text{ ps}$ with a standard deviation of $\sigma = 0.3 \text{ ps}$. In 74% of the tests, the value is within one σ . In 94% of the tests the value is within 2σ . This simulation indicates that tuning is remarkably accurate, facilitating close range beamforming.

In order to evaluate the chip performance, high quality instruments are required. Our equipment is somewhat limited with a 20 Gsample/s scope (Agilent 54855A Infiniium Oscilloscope). However, by using averaging the specification sheet [8] indicates a delta-time accuracy between two edges on a single channel down to 70 fs, provided at least 256 samples are averaged. Based on a number of experiments, we find the practical resolution of the oscilloscope to be approximately 3 ps. Because of these measurement limitations, we have to design our experiments accordingly.

The prototype chip is mounted on a printed circuit board (PCB). The PCB is also equipped with a connector for a microcontroller module hooked up to a PC via USB. The microcontroller is programmed to control the value of the delay-line registers using SPI. The transmitter outputs are routed to eight different SMA connectors. The piggy-back microcontroller is mounted on the right side of the PCB in Fig. 8, while the SMAs are mounted along the edges.

For debugging purposes we have tapped four transmitter trigger signals to separate SMAs (not mounted in Fig. 8). The antenna outputs are connected to the oscilloscope with equal-length cables (1 m) and 50 Ω impedance matching.

The measurements are controlled by MATLAB scripts on a computer interfacing the chip registers through the USB-connected microcontroller and reading results from the oscilloscope using GPIB.

The simplest measurement is to enable each delay element individually and measure the incremental change. The mea-

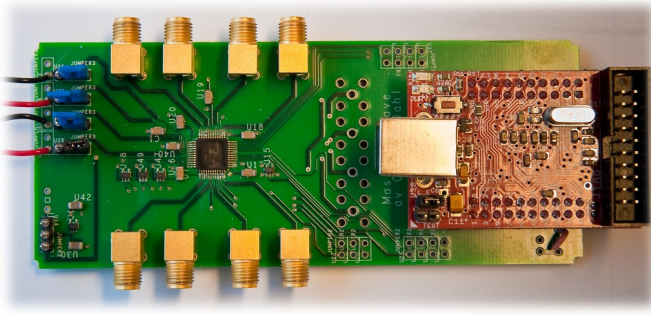


Fig. 8. Photo of PCB made for chip measurements (photo: T. Sæther).

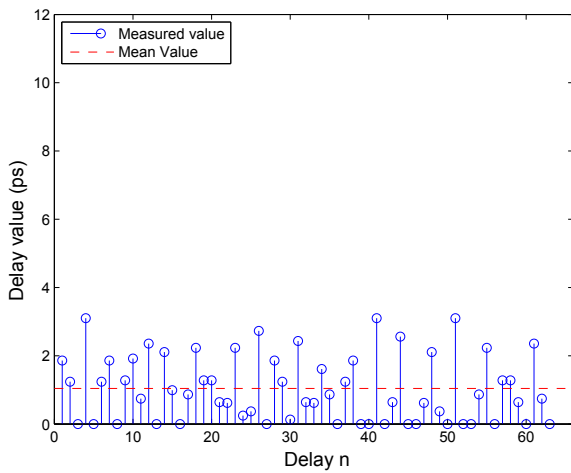


Fig. 9. Measured values for each of the delay-elements in a delay-line.

sured values flip between 0 ps and 3 ps, so each delay-element is measured five times and the average value is recorded (Fig. 9). The average value over all delays is approaching 1 ps.

A better estimation is possible by randomly selecting 40 delay-elements and measure the total delay. By averaging we may estimate the unit delay. Again we do repeated measurements, this time each experiment was repeated 99 times.

The results listed in TABLE I indicate good match with simulations. Largest peak deviation is 3.3 ps which is within the known precision of our oscilloscope. By averaging we estimate the unit delay to be $\tau \approx 1.3$ ps. All these measurements

TABLE I
MEASUREMENTS OF 8 DELAY-LINES ON A PROTOTYPE CHIP, SELECTING
40 RANDOM DELAY ELEMENTS.

Delay-line	# Runs	Mean	Peak deviation	Mean/40
1	99	53.8 ps	2.4 ps	1.3 ps
2	99	44.0 ps	2.8 ps	1.1 ps
3	99	59.6 ps	2.9 ps	1.5 ps
4	99	53.1 ps	3.1 ps	1.3 ps
5	99	56.5 ps	2.9 ps	1.4 ps
6	99	45.3 ps	1.7 ps	1.1 ps
7	99	58.5 ps	2.1 ps	1.5 ps
8	99	47.0 ps	3.3 ps	1.2 ps
Avg	792	52.2 ps	2.7 ps	1.3 ps

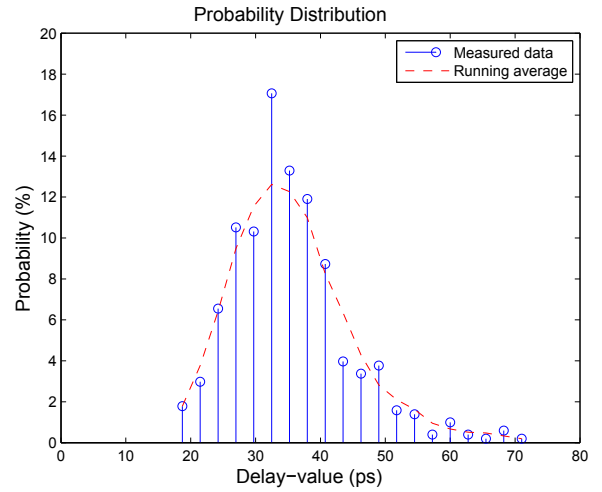


Fig. 10. Measured delay distribution of the eight coarse delay-elements on a single chip.

are from one chip, but results from other chips are similar.

The accuracy of the coarse-tune delay-line is also affecting the overall performance, but can be mended by calibration. These coarse delays are only required for large offsets and are expected to have significant variations in actual delay value.

By measuring each tap of the delay-line and presenting the results as a probability distribution, we are able to analyze the impact of mismatch. These variations are significant, but may to some extent be calibrated by offsetting the fine-tune delay.

By inspecting the probability distribution in Fig. 10, we can see that the delay-values have a somewhat Poisson-like distribution shape. The results show that the steps of the delay-line has a mean value of $\mu = 35.2$ ps, and a standard deviation of $\sigma = 8.9$ ps. Approximately 75% of the steps are within one σ and approximately 96% are within 2σ .

During testing, significant interference between delay-lines were observed. Initially a constant delay value was established in one delay-line (A). Then another delay-line (B) was systematically changed by incrementing the programmed delay. Fig. 11 shows how the presumably constant delay-line A is affected by interference from setup of another delay-line. As delay is increased in delay-line B, we see the presumably constant delay of delay-line A is first increased, then decreased and then increased again. When the firing of transmitter B is done after the firing of transmitter A, no further interference is detected. We observe this effect regardless of physical on-chip distance.

We explain this behavior by rail induced noise. Gates like inverters are rail referred with quite poor power-supply rejection ratio (PSRR). When the pulse transmitters are triggered, rails are strongly loaded, feeding back correlated interference. This systematic interference is showing up as significant change to the delays.

To verify this theory, the pulse generators were turned off by disconnecting the separate rail pads. By measuring the buffered transmitter trigger we find no interference. Based on

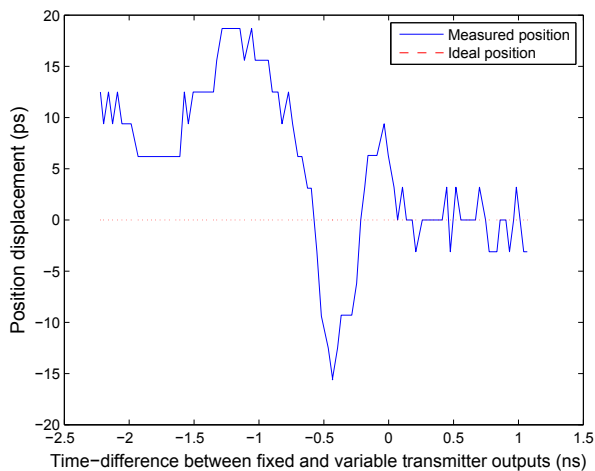


Fig. 11. Interference to the timing of one transmitter caused by another transmitter.

these observations, we see that great care must be taken to prevent cross-talk between pulse generators and delay-lines.

III. CONCLUSION

In this paper, we have presented a novel programmable delay element exploring back-gate tuning. A programmable delay with a relative tuning of 1.1 ps may be used for accurate, close-range beamforming by controlling the triggering

of impulse radio transmitters. Although some inaccuracies and interference is observed, the back-gate tuning is usable for close range beamforming, provided layout is done carefully with appropriate shielding.

ACKNOWLEDGMENT

The authors would like to thank Dept. of Informatics, University of Oslo for supporting our research and for providing instrumentation and chip fabrication. We would also like to thank Novelda AS for their generous contribution of the impulse radio transmitter layout used in this work.

REFERENCES

- [1] [Online]. Available: <http://www.staccatocommunications.com/>
- [2] [Online]. Available: <http://www.novelda.no/>
- [3] R. Merritt, "Report: Ultrawideband dies by 2013." [Online]. Available: <http://www.eetimes.com/showArticle.jhtml?articleID=217201265>
- [4] H. A. Hjortland, D. T. Wisland, T. S. Lande, C. Limbodal, and K. Meisal, "Thresholded samplers for uwb impulse radar," in *Circuits and Systems, 2007. ISCAS 2007. IEEE International Symposium on*, 2007.
- [5] H. A. Hjortland and T. S. Lande, "Cibv integrated impulse radio design for biomedical applications," *IEEE Transactions on Biomedical Circuits and Systems*, vol. 3, no. 2, pp. 13–23, Apr. 2009, to be published.
- [6] T.-S. Chu and H. Hashemi, "A cmos uwb camera with 7x7 simultaneous active pixels," in *Solid-State Circuits Conference, 2008. ISSCC 2008*, 2008, pp. 120–122.
- [7] K. von Arnim, E. Borinski, P. Seegebrecht, H. Fiedler, R. Brederlow, R. Thewes, J. Berthold, and C. Pacha, "Efficiency of body biasing in 90-nm cmos for low-power digital circuits," *Solid-State Circuits, IEEE Journal of*, vol. 40, no. 7, pp. 1549–1556, Jul. 2005.
- [8] [Online]. Available: <http://cp.literature.agilent.com/litweb/pdf/5988-7976EN.pdf>

Paper 14 Electromagnetic impulse radio camera

M. Z. Dooghabadi, T. A. Vu, S. Sudalaiyandi, H. A. Hjortland, T. S. Lande, O. Naess, and S. E. Hamran. “Electromagnetic impulse radio camera”. In: *NORCHIP, 2009*, pp. 1–5, Trondheim, Nov. 2009. doi:10.1109/NORCHIP.2009.5397835. © 2009 IEEE. Reprinted with permission.

My Contributions to This Paper

The beamforming idea using programmable delay elements was elaborated further by a fellow PhD student. Extensive measurements were undertaken to confirm the beamforming idea. I contributed to the design of the delay elements used in the chip which was measured on in this paper. My SPI solution was also used in the chip.

Electromagnetic Impulse Radio Camera

Malihe Zarre Dooghabadi, Tuan Anh Vu, Shanthi Sudalaiyandi, Håkon A. Hjortland,
Tor Sverre “Bassen” Lande, Øivind Næss, and Svein Erik Hamran

Dept. of Informatics, University of Oslo, Norway

Email: malihez@ifi.uio.no

Abstract—An UWB impulse radio programmable transmitter array is presented, usable for electromagnetic beamforming. The linear antenna array may be programmed for focal beamforming using seven impulse radio transmitters. Two different spatial configurations are evaluated for use as beam scanning electronics of an electromagnetic camera. A novel high-precision, programmable delay element is explored for accurate beam control. Measured results shows that the scan line of the beamformer can be steered with sufficient resolution facilitating an electromagnetic CMOS camera.

I. INTRODUCTION

Using electromagnetic waves (EM) for imaging has been around for a long time. The first radar was patented in 1908 and has been refined and developed to a sophisticated device in all sorts of transportation. Another important development is ground penetrating radar (GPR or georadar) also exploring EM. The successful use of EM for longer distance radar and georadar has initiated use for close range imaging like medical diagnosis tools. Some interesting reports on breast cancer diagnosis (Darmstadt) are published [1]. Electromagnetic tomography (EMT) is also under development for short range application. A major challenge in short-range EM imaging is the travelling speed close to light ($3 \cdot 10^8$ m/s), demanding high speed signal processing.

In general the electromagnetic properties of different materials are interesting since EM to some extent shine through materials opaque to light. We may imagine that high-quality EM images of the inside of our body will provide valuable information in medical diagnosis. In addition to existing techniques (X-rays, ultrasound, MR), EM images are indicating electrical properties of the exposed object (permittivity, permeability). Finding cancer tumors in breasts has already been successfully explored with EM [1]. However, great care must be taken to minimize the amount of EM exposure, since some studies are indicating EM may cause cancer [2].

Still, a high-quality electromagnetic imaging device with minimal radiation may be a versatile instrument not only in medical diagnosis. In this paper we will present a potential CMOS implementation of EM camera for short-range application exploring accurate impulse radio beamforming.

A. EMCAM technology

At the Nanoelectronics group at Dept. of Informatics, University of Oslo we are pursuing unconventional design paradigms enabling impulse radio solutions in standard digital

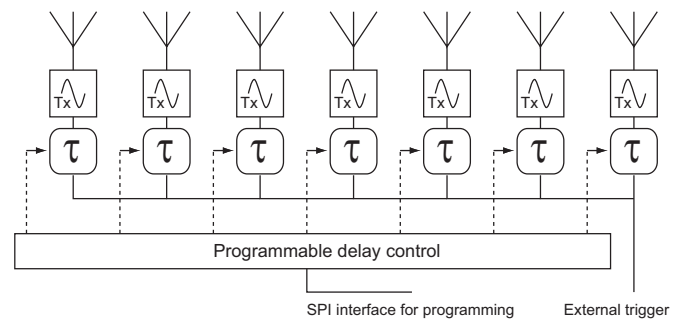


Fig. 1. The transmitter beamforming is achieved by delayed trigger of pulse transmitters.

CMOS operating at microwave frequencies. The Continuous-Time Binary-Value (CTBV) design paradigm has been explored to make a fully functional, single chip impulse radar [3], [4]. A short-range radar has been made with remarkably good results using CTBV design and has since been commercialized [5]. Short-range measurement of backscattered EM energy with millimeter resolution is feasible. However, backscattering from any objects of the exposed field is sensed regardless for location.

Another effort at the Nanoelectronics group is the development of impulse radio beamforming technology using high resolution programmable delay elements for coordinated firing of impulse radio transmitters [6]. Measured performance indicate relative delay programmability in the order of $1 - 2$ ps. When these delay elements are cascaded and inserted in front of impulse radio transmitters, accurate EM beamforming is feasible. As shown in Fig. 1 the impulse radio beamformer is used for relative firing delay of a number of impulse radio transmitters.

Most EM beamformers are found in radar applications and is called phased array antennas. By adjusting the relative phase differences of signals transmitted by several antennas, constructive and destructive interference will set up a radiation pattern a process known as beamforming. Significant literature in this field [7] is available. An interesting difference between impulse based systems and frequency based systems is the lack of grating lobes in impulse radio beamforming.

Reported impulse radio EM beamformers are exploring beamforming in the receivers [8] using phase shifted mixing of backscattered signals.

In this effort we are aiming at a full 3D, short-range

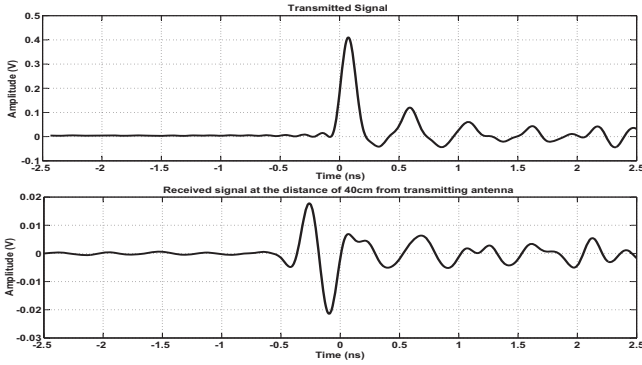


Fig. 2. Transmitted and Received waveforms at the distance of 0.4m from transmitter

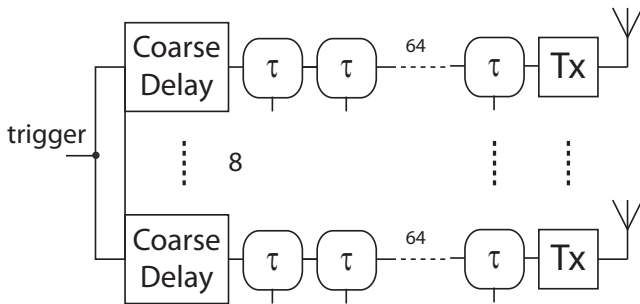


Fig. 3. The fabricated chip consists of eight dual-slope Gaussian pulse transmitters preceded by programmable delay-lines.

EM-camera exploring a two-dimensional transmitting antenna beamformer for directing the beam, while the third dimension (depth) is achieved in the receiver using the high-speed CTBV sampler of the impulse radar. However, several challenges still remains, when implementing such solutions in nanometer CMOS technology.

In the following we will evaluate the implementation reported [6] for extension to a two-dimensional antenna array and explain how impulse radio beamforming and impulse radar technology may be combined to implement a single-chip CMOS EM camera for short-range applications.

II. BEAMFORMER SETUP

A major challenge doing impulse radio solutions in standard CMOS is the low power supply, approaching 1 V in nanometer technology. Even with rail-to-rail operation, it is hard to radiate the permitted energy level set by the US FCC part 15 limit of -41.3 dBm/MHz. Our Gaussian pulse has a peak-to-peak of 0.45 V as shown in Fig. 2.

Initial experiments using low-gain and simple antennas exhibit way too little energy and more antenna gain was required for reliable measurements. The transmitting signal bandwidth is ≈ 4 GHz with a center frequency just above 3 GHz. For increased energy emission, a suitable Vivaldi antenna was developed by another project member (reported in another paper submitted to this conference) with ≈ 10 dB gain.

Measurements reported in [6] indicated significant variations in the delay elements due to transistor mismatch and on-chip interference. Although these variations are expected in CMOS, elaborate calibration was required to compensate.

A. Impulse Radio Beamformer chip

In order to evaluate the quality of delay-line programming a test-chip was made in ST 90 nm CMOS process. Eight identical, edge-triggered impulse radio transmitters for dual-slope Gaussian-shaped pulse transmission were used [6].

As shown in Fig. 3, the firing of each transmitter is controlled by a programmable delay-line consisting of one programmable coarse-tune delay element and 64 programmable fine-tune delay elements. The coarse delay element is made of a chain of 64 dual inverter delays. Each of the delay-outputs are tapped to a multiplexer for control of the delay-value. The multiplexer enables us to choose the number of dual inverter delays to cascade. The unit delay of the dual inverters are approximately 35 ps, which gives us a programming range of 0 to 63×35 ps ≈ 2.2 ns. The firing sequence is triggered by an external trigger for all transmitters and a simple SPI scheme is used for programming the different delay-elements integrated on the chip.

For measurement purposes, one of the transmitters was used as time-reference, so a total of seven transmitters were available for beamforming.

III. EXPECTED PERFORMANCE OF IMPULSE RADIO BEAMFORMING

The expected radiation pattern depends on the array geometry, the elements weights and also the number of elements. The beamwidth depends on the element spacing. The larger element spacing the narrower beamwidth and the better resolution. The number of elements determines the number of sidelobes and the peak of the main lobe. UWB beamforming is time and space signal processing [9], the array pattern for a linear receiving array of N elements with equal element spacing of d is:

$$y(\Phi, \Phi_0, t) = \frac{1}{\sum h_n} \sum_{n=0}^{N-1} h_n s(t + n(\tau_\Phi - \tau_{\Phi_0})) \quad (1)$$

where $s(t)$ and h_n are the impulse signal and element weighting coefficient, respectively. If the signal propagation speed is c and angle of incident signal is ϕ , then the signal with respect to the first sensor reaches the n_{th} element after $n\tau_\phi$, where $\tau_\phi = \frac{d}{c} \sin(\phi)$. For the steering angle of ϕ_0 , the steering delay is $\tau_{\phi_0} = \frac{d}{c} \sin(\phi_0)$. In Fig. 4 the expected lobe pattern of 7 elements receiving array is shown. Unlike frequency based beamforming no grating lobes are present.

A. Measured Performance

For these measurements, a uniform linear transmitting array with 7 elements is used. In Fig. 2 the transmitted waveform from one transmitter and received waveform at the distance of 40 cm from the transmitting antenna are presented. According

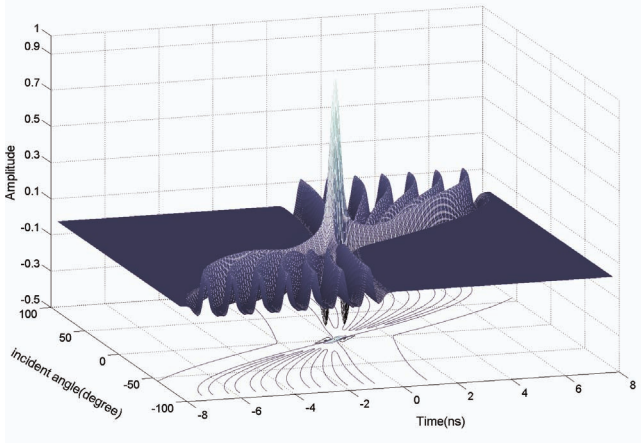


Fig. 4. Modelled beamforming with seven receivers.

to Friis equation if we have two transmitting and receiving antennas, the received power to transmitted power ratio is:

$$\frac{p_r}{p_t} = G_r G_t \left(\frac{\lambda}{4\pi R} \right)^2 \quad (2)$$

where G_t and G_r are transmitting and receiving antenna gain, respectively. λ is wavelength and R is distance. So, as we goes far away from transmitting antenna, the received signal attenuated. This effect is shown in Fig. 2 in which the received signal at the distance of 0.4m is weaker than the transmitted signal.

B. Beamformer

In a linear transmitting array, for focusing on different points, transmitters should be delayed in a way that on the focal point the transmitted waves reach simultaneously. In the following, the element delay value of a uniform linear array is calculated. In an array of N elements as shown in Fig. 5 the propagation time of transmitted waves to reach on focal point X , are:

$$\begin{aligned} t_n &= \frac{D_n}{c}, \quad t_{n-1} = t_{n+1} = \frac{\sqrt{D_n^2 + d^2}}{c}, \\ t_{n-2} &= t_{n+2} = \frac{\sqrt{D_n^2 + (2d)^2}}{c}, \dots \end{aligned} \quad (3)$$

where c is the propagation speed and equals to $3 \cdot 10^8$ m/s. If $t_{\max} = \max(\dots, t_{n-2}, t_{n-1}, t_n, t_{n+1}, t_{n+2}, \dots)$, then the delay value of elements are:

$$\tau_n = t_{\max} - t_n, \tau_{n-1} = \tau_{n+1} = t_{\max} - t_{n-1}, \dots \quad (4)$$

Transmitters are delayed based on the equation (4).

1) *Calibration and measurement setup:* All transmitters must be calibrated to account for transistor mismatch. As reported in [6], CMOS production mismatch introduces major deviations in gate delay. To some extend we were able to align seven transmitters by using one transmitter as a reference. However, this simple scheme is only adjusted for all transmitters in phase. When varying the delays for focusing, significant

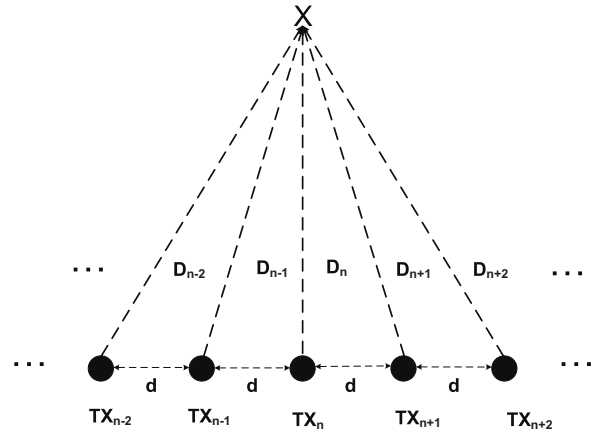


Fig. 5. Uniform transmitting linear array

errors were introduced. In order to ensure correct delays, each actual delay setting was measured using an oscilloscope before beamforming experiments. For future development, calibration procedures must be incorporated as a chip function.

Another limitation is our small lab facilities without any absorbers. Significant reflections and external noise is degrading our measurements, but even so, our results are quite promising.

The experiments are done in time domain using a receiving antenna connected to an oscilloscope and directly measuring signal strengths with build-in scope functions.

2) *Experimental results:* To see the effect of antenna spacing on the EM pattern, a 7 elements uniform linear array with antenna spacing of 10 cm which is about one wavelength and 5 cm which is half a wavelength, is used. Fig. 6 shows the measured EM field of the array with antenna spacing of 5 cm, focused at (20 cm, 50 cm). The EM field of array with antenna spacing of 10 cm is shown in Fig. 7 with a focal point at (30 cm, 50 cm). In these figures the EM field in light areas is stronger than the dark areas. The X axis indicate the transmitter line-up and the Y axis gives the depth. The EM field is measured form 25 cm in the Y direction.

Figures 8 and 9 shows the array pattern of the array with 5 cm element spacing and 10 cm element spacing at the depth of 50 cm, respectively. As shown in figures 6 and 8, because of interferences the measured focal point is offset to (22.5 cm, 50 cm).

There are different definitions for calculating the resolution [10]. In the paper we adopt the -3 dB attenuation point as the beam limit or resolution. The -3 dB resolution in Fig. 8 is about 11.5 cm, and in Fig. 9 is about 9.5 cm which is about one wavelength (the center frequency is 3.2 GHz and the wavelength is 9.4 cm). So, the larger antenna spacing, the narrower mainlobe and the better resolution.

The 7 elements transmitting array with element spacing of 10 cm is programmed to focus on (60 cm, 50 cm). The measured EM field is shown in Fig. 10. The focal point is on (55 cm, 50 cm) with a steering angle about 26.5°. The radiation pattern at depth of 50 cm is shown in Fig. 11, resolution is

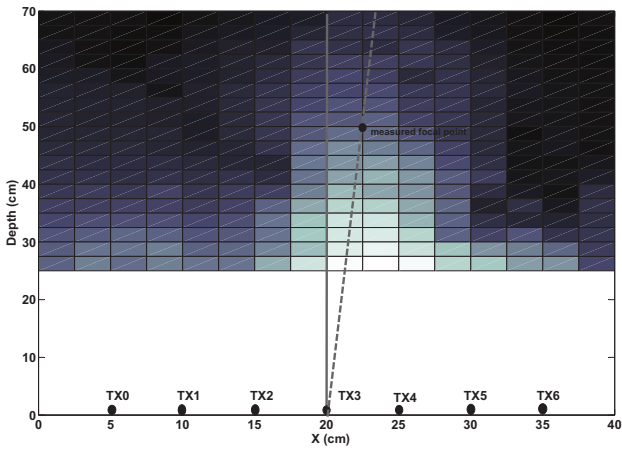


Fig. 6. EM Field of 7 TXs with antenna spacing of 5 cm, Focal point at (22.5cm,50cm)

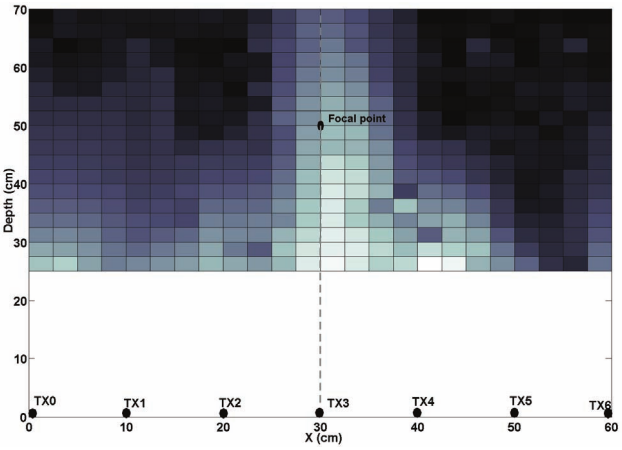


Fig. 7. EM Field of 7 TXs with antenna spacing of 10 cm, Focal point at (30cm,50cm)

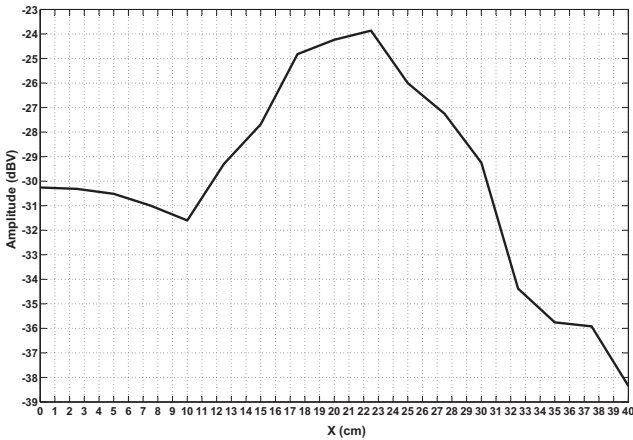


Fig. 8. Radiation pattern of 7 elements array with 5cm spacing at the depth of 50 cm

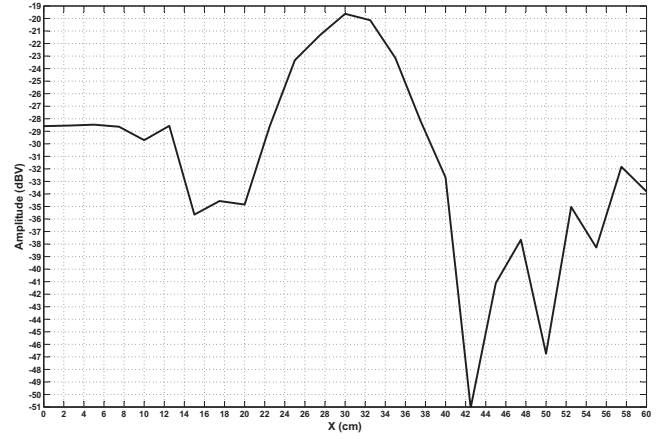


Fig. 9. Radiation pattern of 7 elements array with 10cm spacing at the depth of 50 cm

about 11cm.

There are some sidelobes in radiation pattern of figure 9 and 11. the sidelobes in Fig. 9 are approximately at 6cm, 47cm and 57cm which correspond to -25.6° , 17° and 28.7° and in Fig. 11 are approximately at 6cm, 20 cm and 37.5 cm that correspond to -25.6° , 11.3° and 8.5° , respectively. We blame these sidelobes mainly to inaccurate calibration and interference.

In narrow band arrays with element spacing larger than half a wavelength, grating lobes appear. These grating lobes may approach the signal strength of the main lobe and limits antenna spacing. But, in impulse radio arrays the antenna spacing may be larger without any aliasing or grating lobes. So, there is not any grating lobes in figures 9 and 11, while the element spacing is about one wavelength.

In general we find the experimental results very promising, but will require further investigation for a two-dimensional structure. The challenge of tuning delays within a few ps is feasible, but require calibration.

IV. CONCLUSION

We have experimentally shown that short-range electromagnetic beamforming requiring picosecond delay tuning is feasible in standard CMOS technology. The measurement results shows that the beam can be controlled and focused with a spatial resolution of about one wavelength. Also the effect of element spacing is investigated, for reduced focal size, the antenna element spacing should be increased, which is feasible with impulse radio transmitters. Based on these promising experimental results, we are pursuing solutions for making a CMOS EM camera using transmitting beamforming in two dimensions, while exploring CTBV impulse radar technology in the depth dimension.

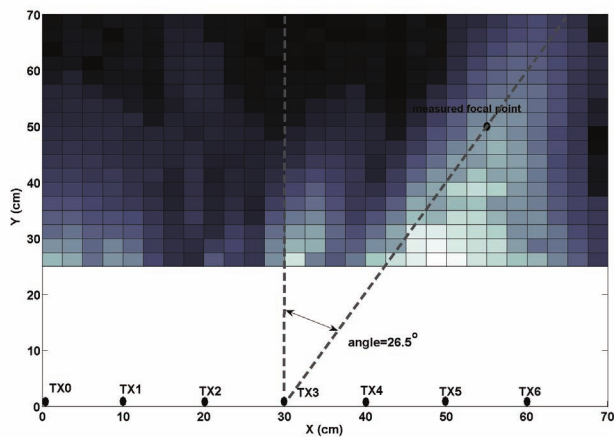


Fig. 10. EM Field of 7 TXs with antenna spacing of 10 cm, Focal point at (55cm,50cm)

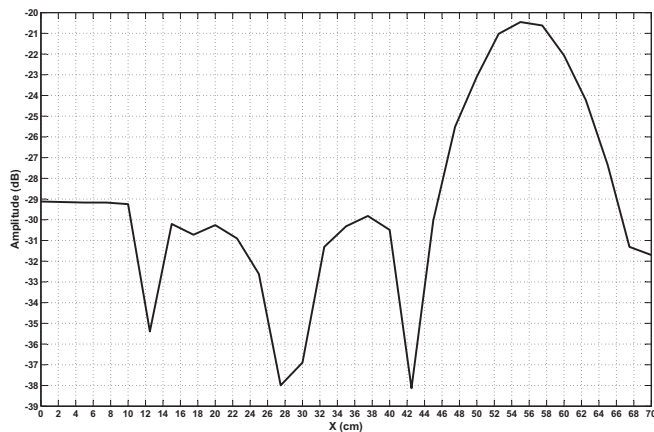


Fig. 11. Steered radiation pattern of 7 elements array with 10cm spacing at the depth of 50 cm

ACKNOWLEDGMENT

The authors would like to thank Dept. of Informatics, University of Oslo for supporting our research and for providing instrumentation and chip fabrication. We would also like to thank Novelda AS for their generous contribution of the impulse radio transmitter layout used in this work.

REFERENCES

- [1] E. J. Bond, X. Li, S. C. Hagness, and B. D. V. Veen, "Microwave imaging via space-time beamforming for early detection of breast cancer," *Signal Processing*, vol. 51, no. 8, pp. 2909–2912, Aug. 2003.
- [2] E. R. Schoenfeld, E. S. O'Leary¹, K. Henderson¹, R. Grimson, G. C. Kabat, S. Ahnn, W. T. Kaune, M. D. Gammon, and M. C. Leske, "Electromagnetic fields and breast cancer on long island: A case-control study," *American Journal of Epidemiology*, vol. 158, no. 1, pp. 47–58, 2003.
- [3] H. A. Hjortland, D. T. Wisland, T. S. Lande, C. Limbodal, and K. Meisal, "Thresholded samplers for uwb impulse radar," in *Circuits and Systems, 2007. ISCAS 2007. IEEE International Symposium on*, 2007.
- [4] H. A. Hjortland and T. S. Lande, "Ctby integrated impulse radio design for biomedical applications," *IEEE Transactions on Biomedical Circuits and Systems*, vol. 3, no. 2, pp. 13–23, Apr. 2009.
- [5] [Online]. Available: <http://www.novelda.no/>
- [6] Ø. Dahl, H. A. Hjortland, T. S. Lande, and D. T. Wisland, "Close range impulse radio beamformers," in *To be published at: 2009 IEEE International Conference on Ultra-Wideband, September 9th to 11th 2009, Vancouver, Canada, 2009*.
- [7] J. R. Guerri, "Close range impulse radio beamformers," in *Space-time adaptive processing for radar*, 2003.
- [8] T.-S. Chu and H. Hashemi, "A cmos uwb camera with 7x7 simultaneous active pixels," in *Solid-State Circuits Conference, 2008. ISSCC 2008*, 2008, pp. 120–122.
- [9] S. Ries and T. Kaiser, "Ultra wideband impulse beamforming: it is a different world," *Signal Processing*, vol. 86, no. 9, pp. 2198–2207, Sep. 2006.
- [10] D. H. Johnson and D. E. Dudgeon, "Array signal processing-concepts and techniques," in *Prentice Hall*, 1993.

Paper 15 High precision calibrated digital delay element

M. Z. Dooghabadi, H. A. Hjortland, and T. S. Lande. “High precision calibrated digital delay element”. *Electronics Letters*, Vol. 47, No. 9, pp. 564–565, Apr. 28, 2011. doi : 10.1049/el.2011.0395. © 2011 IET. Reprinted with permission.

My Contributions to This Paper

This paper presents a calibration circuit for tunable delay elements which reaches a measuring precision of one picosecond. The solution can be used for IR-UWB radar beamforming, which requires such high-precision tunable delay elements. I came up with most of the design for the calibration circuit, and my SPI and microcontroller solutions were used for interfacing. I also helped review the paper.

High precision calibrated digital delay element

M. Zarre Dooghabadi, H.A. Hjortland and T.S. Lande

A novel high precision digital delay element for use in impulse radar beamforming is presented. By using digitally tunable delay elements, a simple digital calibration procedure is possible using a novel, non-inverting ring oscillator in combination with an accurate clock reference. All process dependent errors, such as mismatch and process variations, are corrected and may be stored as a digital calibration pattern.

Introduction: In modern narrowband radar technology, phased-array solutions are used for directional, electronic steering of radar beams, avoiding mechanical steering. Recently, the original impulse radar has been revived [1, 2], exploring high processing gain and low-power impulse emission. With the unlicensed ultra-wideband (UWB) spectrum released by FCC (3.1–10.6 GHz), impulse radio solutions are becoming even more attractive. An advantage of impulse radar systems is improved depth resolution due to the emission of short and thus wideband pulses. This has made short-range radars with millimetre depth resolution possible. For sufficient processing speed, a novel processing paradigm is utilised, known as continuous-time value (CTBV) [2]. A similar approach was also proposed in [3, 4], named CTDA. The CTBV processing technique facilitated the design of a single-chip CMOS radar, now even commercialised [5].

As with narrowband systems, impulse radar systems may also be used for beamforming. By utilising the co-ordinated transmission of pulses, constructive interference appears and directional electromagnetic beams may be radiated. An interesting advantage of impulse radar transmission beamforming is the absence of grating lobes regardless of antenna spacing, thus avoiding directional ambiguity.

To have an accurate co-ordination of pulse firing, high-resolution delay adjustments of the order of a few picoseconds are required. For example, in a linear UWB transmitting array with seven elements and antenna spacing of 10 cm at a centre frequency of 6.85 GHz with -3 dB beam resolution of one wavelength at the depth of 50 cm (which is a beam resolution of about 5°), a minimum temporal resolution of 5–10 ps is required. By programming the delay line with such a temporal resolution, an angular beam control of approximately 5° is feasible.

In a typical nanometre technology, the gate delay of an inverter is approximately 10 ps. Adding production spread of up to 20% due to mismatch, digital gate delays do not provide sufficient resolution nor precision. We have previously published a simple, digital tuning technique using back-gate biasing [6, 7] enabling digital tuning of single inverters by a few picoseconds. In this Letter, we characterise delay elements by using a ring oscillator-based calibration circuit with a precision of a few picoseconds to compensate for process variations and device mismatch. It is, however, still vulnerable to dynamic noise, e.g. jitter, and rapid temperature variations.

Calibration circuit: The proposed calibration circuit consists of three parts, the calibration loop, the reference pulse generator, and the loop counter, as indicated by dashed rectangles in Fig. 1. A digitally programmable fine-tuning delay line with 65 settings facilitates adjustment up to approximately 500 ps in steps of 7–8 ps [6, 7]. It has an absolute delay of about 2 ns, but this is not a problem since the same delay is added to all the Tx-antennas in the array. The adjustable delay line is combined with a static delay of 8 ns, giving a total delay of approximately 10 ns, which matches our expected pulse repetition frequency (PRF). The two delay lines may be configured in a loop, which we here call the calibration loop, in order to measure their actual delay. The calibration loop is basically a non-inverting ring oscillator where we send a pulse around the loop rather than an edge, like in a normal ring oscillator. The period of the non-inverting ring oscillator will only be dependent on the rising edge propagation delay through the loop, rather than the sum of the rising and falling edge propagation delays. This is exactly what we are after, since we want to calibrate the propagation delay of a rising edge through the delay element. The period of the rising edge can be measured by counting the number of oscillations for an accurately controlled period of time. The actual period of the ring oscillator is then simply the reference time interval divided by the loop count. A timing diagram is shown in Fig. 1.

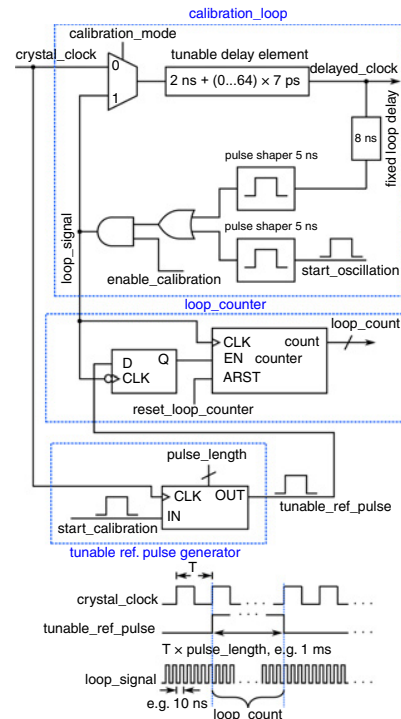


Fig. 1 Block diagram of proposed calibration circuit

For controlled operation of the oscillator, the loop is closed by the 'calibration_mode' signal, as shown in Fig. 1. First, the 'enable_oscillation' signal is set to '0' to stop any oscillation in the loop, then it is set to '1' to enable the loop. After that, a trigger signal on the 'start_oscillation' input starts the actual oscillation. Maintaining the loop oscillation is done by adding a pulse shaper. The pulse shaper is a one shot circuit, which can accept any input signal with pulse width shorter or longer than 5 ns, and it generates an output pulse with constant width of 5 ns. The tunable reference pulse circuit generates a reference pulse with a known length based on an input reference clock only required for calibration. As described, it is used to enable counting of the ring oscillator output in the loop counter for a precise time interval. The width of the reference pulse is set by a programmable 'pulse_length' register.

When a measurement is initiated, the actual delay including static and dynamic errors will be averaged by the counter. If we want a loop delay of say 10 ns, a 1 ms reference pulse should give a counter value of approximately 100000 if the delay is correctly programmed. By repeatedly measuring and tuning the delays, an accurate calibration is found. The accuracy is in part determined by the amount of averaging, i.e. the width of the reference pulse. If we want to increase the delay by 100 ps then we repeat this procedure with a target loop delay of 10.1 ns. This will increase the delay value by 100 ps.

Measurement results: To evaluate the performance of the proposed calibration circuit, an implementation in TSMC 90 nm CMOS technology was made. The programmable delay line is a cascade of 64 dual-inverter buffers. In each of these dual inverters, the bulk voltage of the PMOS transistor in one of the inverters is switched between two values in order to gain the fine delay tuning [6, 7].

All control signals are generated by an on-chip SPI controller. A microcontroller is used for SPI programming and clock signal generation. The microcontroller itself is programmed through a USB interface.

A necessary condition for correct calibration is an accurate clock reference. If we have a crystal clock with 100 ppm precision, a 1 ms reference pulse will give a maximum variation of about 0.1 μ s. This variation, when divided by the loop counter value, will be very small, e.g. for the loop count of 100000, the maximum error is 1 ps, which is sufficient. It should be noted that also dynamic errors like jitter noise, rail-noise and other noise sources are all averaged in this calibration procedure. Therefore dynamic errors are reduced inversely proportional to the reference pulse width.

For the following measurement results, the reference pulse width is set to 1 ms. In the programmable delay line there are 64 delay steps. We sweep

this setting and measure the total delay. We use 100 runs per delay step to average out and to quantify the noise in the measurement results.

Fig. 2a shows the increase in the absolute loop delay as the delay is swept. The irregularities apparent for settings of 0, 16, 32, 48, and 64 are due to imperfect layout with folding and different loading. The measured delay step for each programmable delay element is shown in Fig. 2b. All the delay steps are 7 to 8 ps except for fine delay settings of 0, 16, 32, 48, and 64, which are due to the layout issue.

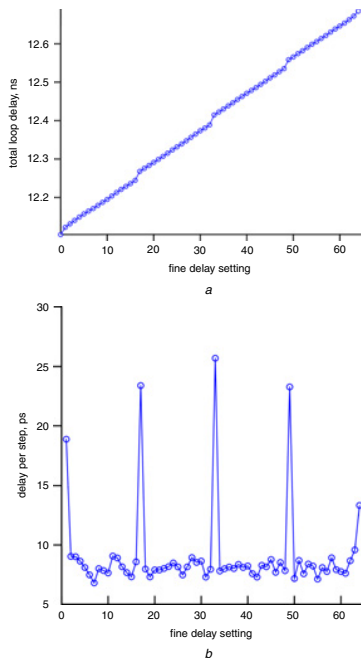


Fig. 2 Measured absolute loop delay and measured delay steps of programmable delay line for 65 delay settings

a Measured absolute loop delay for 65 delay settings

b Measured delay steps of programmable delay line for 65 delay settings

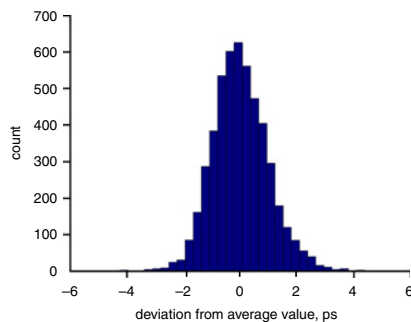


Fig. 3 Variation in measured loop delay from run to run for one delay setting

To show the reliability of the calibration circuit, repeated measurements were done with a fixed programmed delay value, displayed in Fig. 3. The results are quite convincing, giving an RMS value of 1 ps. To reduce RMS error a reference pulse with longer width can be used, but here we are using the calibration circuit to characterise the delay lines for beamforming and this accuracy is sufficient for our purpose.

Conclusion: A novel digitally programmable delay element with tuning steps of 7–8 ps can be calibrated with a measuring precision of 1 ps RMS error using a reference clock. Measured results in 90 nm technology are provided, indicating excellent performance. Although dynamic deviations are not corrected, we anticipate that the calibrated delay element may be usable for other applications as well, such as simple on-chip clock generation.

© The Institution of Engineering and Technology 2011

12 February 2011

doi: 10.1049/el.2011.0395

One or more of the Figures in this Letter are available in colour online.

M. Zarre Dooghabadi, H.A. Hjortland and T.S. Lande (*Department of Informatics, University of Oslo, Oslo, Norway*)

E-mail: malihezd@ifi.uio.no

References

- 1 McEwan, T.E.: 'Ultra-wideband radar motion sensor'. U.S. Patent 5 361 070, 1 Nov. 1994
- 2 Hjortland, H.A., and Lande, T.S.: 'CTBV integrated impulse radio design for biomedical application', *IEEE Trans. Biomed. Circuits Syst.*, April 2009, Vol. 3, pp. 79–88
- 3 Tsividis, Y.P.: 'Continuous-time digital signal processing', *Electron. Lett.*, 2003, **39**, pp. 1551–1552
- 4 Li, Y.W., Shepard, K.L., and Tsividis, Y.P.: 'Continuous-time digital signal processors'. 11th IEEE Int. Symp. on Asynchronous Circuits and Systems, New York, NY, USA, 2005, pp. 138–143
- 5 www.novelda.no
- 6 Dahl, O., Hjortland, H.A., Lande, T.S., and Wisland, D.T.: 'Close range impulse radio beamformers'. IEEE Int. Conf. on Ultra-Wideband, Vancouver, Canada, 2009, pp. 205–209
- 7 Zarre Dooghabadi, M., Vu, A.T., Sudalaiyandi, S., Hjortland, H.A., Lande, T.S., Næss, Ø., and Hamran, S.E.: 'Electromagnetic impulse radio camera'. IEEE Conf. Proc. of 27th NORCHIP, Trondheim, Norway, November 2009

Paper 16 An ultra-wideband receiving antenna array

M. Z. Dooghabadi, H. A. Hjortland, and T. S. B. Lande. “An ultra-wideband receiving antenna array”. In: *Circuits and Systems (ISCAS), 2013 IEEE International Symposium on*, pp. 2373–2376, Beijing, China, May 2013. doi:10.1109/ISCAS.2013.6572355. © 2013 IEEE. Reprinted with permission.

My Contributions to This Paper

This paper presents a radar with receive-beamforming. The system is based on the radar design which we present in other papers and which I have come up with much of the design for. I was not much involved in the design of the quantizer and pulse generator implementations used on this chip, though. I came up with most of the design for the calibration circuit, which is used for the delay-and-sum beamforming here. I also took part in discussions about how to do the beamforming, my SPI and microcontroller solutions are used for interfacing with the chips, and I helped review the paper.

An Ultra-Wideband Receiving Antenna Array

Malihe Zarre Dooghabadi, Håkon A. Hjortland, and Tor Sverre “Bassen” Lande
Department of Informatics, University of Oslo, Norway, Email: malihez@ifi.uio.no

Abstract—An UWB impulse radio receiving antenna array based on delay and sum beamforming is presented. The array is implemented using a simple impulse radio radar transceiver. The presented radar transceiver is fabricated in 90 nm TSMC technology, and it has a high speed sampler with sampling rate of 10 GHz which recovers the received signal with sufficient resolution. Using a programmable delay element with medium and coarse settings, an electronic beam steering step of about 10-15° is feasible. A high precision calibration circuit is explored for delay adjustment with high resolution. The measurement results show that the array with antenna spacing of about one wavelength has a cross-range resolution of 17 cm, which is about one and half of the wavelength.

I. INTRODUCTION

The ultra-wideband (UWB) imaging array can be used in several applications such as medical imaging and diagnosis, health-care and security scanning. The UWB antenna arrays may be categorized in two main groups, delay and sum beamforming arrays and spatial diversity arrays [1]. The spatial diversity receiving array is used to improve the signal and image quality in a multipath fading environment. The antenna elements should be widely spaced so that the received signals at each of the antenna elements are uncorrelated. In delay and sum beamforming arrays the transmitted or received signals are delayed to have constructive interference at the focal point or the desired direction. So, the antenna arrays improves the directivity of the main beam, and also provides electronics steering without mechanical rotation.

In conventional narrowband beamforming arrays, the time delay is approximated with a constant phase shift, but the impulse radio UWB arrays are True-Time-Delay (TTD)-based. In [2], an on-chip coplanar strip line and tapped delay trombone line are used for delay adjustment, but this architecture requires a large number of on-chip inductors, transmission lines, and complex trombone tap circuit structures. In [3] the beamformer architecture is based on the double quadrature receiver in which the received impulse signal is down converted to the baseband for signal detection, and a baseband digital phase shifter is used as a TTD element.

In this paper a linear receiving array is presented which is implemented with several IR-UWB radar transceivers based on the Continuous-Time Binary-Value (CTBV) technique. The CTBV technique is based on the inherent, process-dependent gate delays, which avoids power-demanding clocked schemes. In CTBV signal processing domain the signal amplitude is binary value of either 1 or 0, while it is continuous in time domain [4].

Using CTBV technique, the presented IR-UWB transceiver has a simple circuit structure, which is mostly realized with

digital circuits. The IR-UWB transceiver has a high speed sampler with sampling rate of about 10 GHz that facilitates recovering the back scattered signal with sufficient resolution. By using a programmable delay element with medium tuning steps of 45 ps and coarse tuning steps of 900 ps, an electronic beam steering step of about 10 – 15° is feasible.

To have an accurate co-ordination of received signals high-resolution delay adjustments of the order of a few tens picoseconds are required. Using an on-chip calibration circuit and exploiting a calibration procedure all the process variations and mismatches in the transceiver circuits, and also the mismatches between PCB boards and cables can be compensated.

The paper is organized as follows. II gives an overview of delay and sum beamforming and antenna array design, and III explains the presented antenna array system and IR-UWB radar transceiver. The measurement results are presented in IV and V concludes the paper.

II. UWB BEAMFORMING AND DESIGN CONSIDERATIONS

In a receiving array, signals from one back-scattering object will reach the antennas at different times proportional to the time-of-flight distance. By compensating these time delay differences and signal alignment, constructive interference and beam-steering at the desired direction, and destructive interferences at other directions can be obtained. The output of the delay and sum beamforming array is: $y(t) = \sum_{i=0}^{n-1} w_i s_i(t - \Delta_i)$ [5], where $s_i(t)$ and w_i are impulse signal and element weighting coefficient, respectively. n is number of antenna elements, and Δ_i is the steering delay.

In an antenna array the beamwidth is inversely proportional to the aperture size. So, with the same number of antenna elements but larger antenna spacing the beamwidth gets narrower and hence the cross-range resolution improves. However in narrowband antenna array with antenna spacing larger than a one-half of the wavelength grating lobes appear which results to spatial aliasing and ambiguity in direction. In UWB beamforming due to the short duration of the emitted impulse signals, there is less interference outside the focusing location, therefore the level of the grating lobes is decreased. Although larger inter-element spacing provides better resolution, but antenna arrays with smaller antenna spacing have better SNR.

The ideal sidelobes level (ISL) is ratio of the maximum level of the sidelobe to the mainlobe level, and it measures the array ability to reject the noise and unwanted signal. ISL is defined as: $ISL = 20 \log_{10} \frac{1}{n}$ [6], where n is the number of antenna elements. So, within the same aperture size increasing the number antenna elements results in lower sidelobes level, and higher dynamic range.

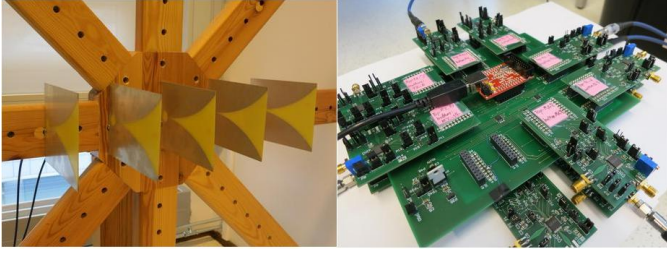


Fig. 1. Antenna array and beamformer system.

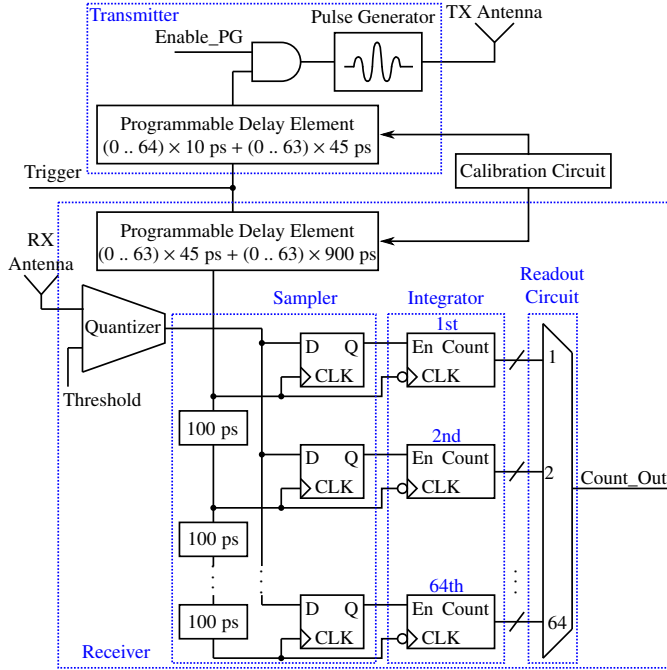


Fig. 2. Transceiver architecture.

III. LINEAR UWB RECEIVING ARRAY

A. Overview of the Array System

The presented linear receiving array consists of several single chip IR-UWB radar transceivers which are used as the receiving antenna elements. A radar transceiver is also used as a point source radiator and transmitter. For energy emission and receiving, suitable Vivaldi antennas with $\approx 10\text{dBi}$ gain are used. Fig. 1 shows the array setup. As shown in Fig. 1 several radar transceiver modules are mounted on the main beam-forming PCB board. A micro-controller is used on the main PCB board for the signal controlling and SPI programming of all the radar transceivers. The micro-controller is programmed through USB interface.

B. Transceiver Architecture

The block diagram of the IR-UWB radar transceiver is shown in Fig. 2. The transceiver consists of several sub circuit blocks of transmitter, receiver, calibration circuit, and SPI module for generating all the controlling signals.

1) *Transmitter*: The transmitter is simple and all-digital, and it has two sub circuits of a programmable delay element and a pulse generator. The programmable delay element is

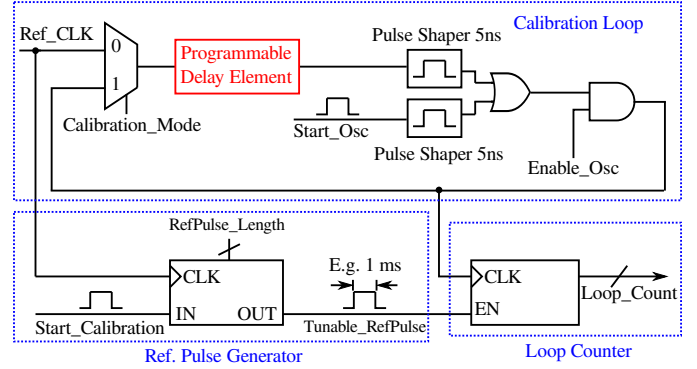


Fig. 3. Calibration circuit (In transceiver the Trigger is used as Ref_CLK).

implemented with two sub delay elements of fine and medium delay lines. The fine delay line is a cascade of 64 dual-inverter buffers in which the back-gate well voltage of the PMOS transistor in one of the inverters can be changed, giving a programmable delay element with 10 ps step size. The medium delay line is a tapped chain of 63 dual-inverter buffers with delay values of 45 ps.

The presented pulse generator in [7] is used to generate the high order Gaussian pulses. Several simple pulses construct a first order Gaussian envelope. The bandwidth is determined by the envelope, and the center frequency is determined by the frequency of the generated pulses. By setting *Enable_PG* to 0, the pulse generator will be disabled, so that it does not introduce any interference or induced rail noise during receive mode.

2) *Receiver*: The receiver circuit operates in CTBV domain, and it is implemented with a programmable delay element, a 1-bit quantizer, a continuous time sampler, a digital integrator, and a readout circuit. The programmable delay element inside receiver is used for ranging and consists of medium and coarse delay lines. The medium delay line has step size of 45 ps, and the coarse delay line consists of a tapped chain of 63 delay elements, each of which is 20 dual-inverter buffers, making a step size of 900 ps. The programmable delay element with mentioned specifications provides sufficient delay value for close range imaging and compensating all the delays which are due to the cables length.

The single ended 1-bit quantizer at the receiver front end compares the incoming signal with the threshold voltage and gives a single bit digital output of either 1 or 0. The threshold voltage level can be set from 0 to VDD [8]. The D flip-flops are used as sampling elements of the continuous time sampler. The sampler is implemented with a delay line which is tapped by the D flip-flops in the steps of 100 ps, and hence facilitates a sampling rate of 10 GHz.

After the quantized output is sampled, the counters in the integrator which are connected to the D flip-flops increment if the 1 is sampled. This integration procedure is repeated while the threshold voltage is swept, as a result those counters which are clocked at the high peaks of the received signal will count more, and those counters which are clocked at the low peaks of the received signal will count less. By integration

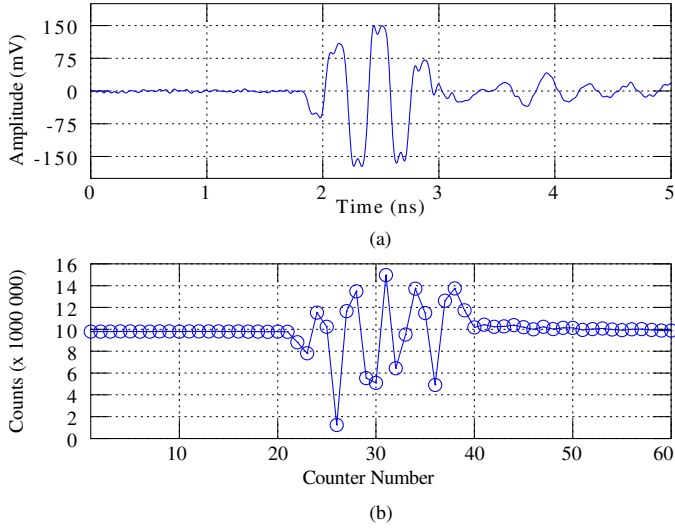


Fig. 4. (a) The measured Gaussian pulse. (b) The sampler readout of the receiving array after calibration.

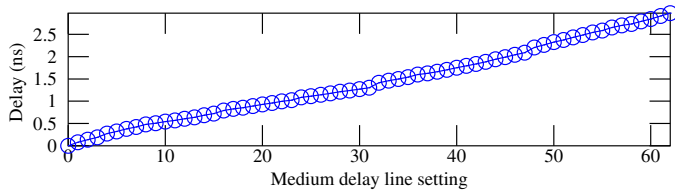


Fig. 5. Measured delay value of the medium delay line for 64 delay settings.

and sweeping the threshold voltage the received signal can be recovered with high resolution.

3) *Calibration Circuit*: One of the important parameter in delay and sum beamforming is adjusting the delay value with high resolution in order to have accurate signals alignment at the desired direction. Due to the process variation and production mismatch, there are some deviation in the gate delays. Therefore a calibration scheme which has been presented in [9] is used to compensate for the process variations and mismatch.

The conceptual block diagram of the calibration circuit is shown in Fig. 3. In calibration mode, by setting *Enable_Osc* to high and sending a pulse on the *Start_Osc* the calibration loop which basically is a ring oscillator, will be enabled and starts oscillating. The oscillation should be kept for a little while to be stabilized. Calibration can be started by sending a pulse on the *Start_Calibration*, then a tunable reference pulse will be generated. As long as the reference pulse is high the loop counter counts the oscillation cycles. The length of the reference pulse can be set using the *RefPulse_Length*.

The period of the loop signal is the reference pulse divided by the loop count value, e.g. for the reference pulse of 1 ms, and loop count of 100 000, the period of the loop signal is 10 ns. Thus, if we want to decrease or increase the delay by 100 ps, we repeat this procedure to have a loop signal with period of 9.9 ns or 10.1 ns.

IV. MEASUREMENT RESULTS

The uniform receiving array is implemented with IR-UWB radar transceivers for near-field beamforming. The presented

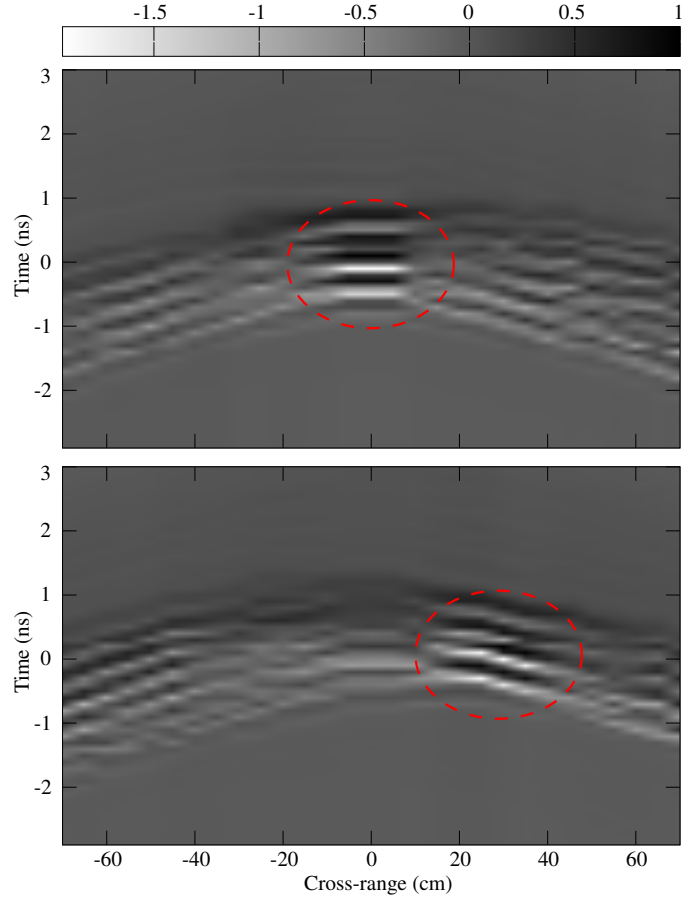


Fig. 6. The top graph is the EM radiation which measured by the array when the radiator is on the array center line. The 2nd graph is the measured EM radiation by the array when the radiator is 30 cm off the array center line.

transceiver is fabricated in TSMC 90 nm CMOS low power process.

In the measurement system setup, the array consists of 5 receiving antenna elements with antenna spacing of 10 cm which is about one wavelength. For measurement purpose, one more radar transceiver is used as a transmitter and a point source for energy emission, which is placed in front of the receiving array.

The emitted signal is a high order Gaussian pulse with center frequency of 2.5 GHz as shown in Fig. 4(a). As shown in Fig. 2 the sampler has 64 sampling elements which tap the delay line in steps of 100 ps, so the sampling frame is about 6.4 ns. In the presented measurement results, the data from the first 60 sampling elements are used.

A. Antenna Array Calibration

In delay and sum beamforming before varying the delay of the antenna elements for beamsteering and focusing, all the receiving elements must be calibrated to account for transistor mismatches between the chips, and also compensate all the mismatches of the PCB boards and cables.

For array calibration, the transmitter is placed in front and on the center line of the receiving array. Then, the programmable delay element of each of the receivers is tuned to

have the received signal in the middle of the sampling frame. Using this procedures all the receiving antenna elements can be aligned and calibrated precisely.

The readout data from calibrated receivers are summed together and presented in Fig. 4(b). As can be seen in Fig. 4(b), after array calibration the received signal is in the middle of the sampling frame.

B. Delay Line Calibration

After the antenna array is calibrated to compensate all the above mentioned mismatches, the receivers can be delayed according to the required delay value for beamsteering at the desired direction. Using the presented calibration circuit in III, an accurate delay adjustment which is important in electronic beamsteering can be achieved.

In order to calibrate the programmable delay elements and set them to the desired value, a reference signal (Trigger signal in transceiver) with clock frequency of 3 MHz is used and the tunable reference pulse signal is set to 1 ms. Then, the delay setting of the medium and/or coarse delay lines are swept to get the required delay increase or decrease with high precision of few pico second RMS error [9].

For illustration of the calibration circuit functionality, the calibration measurement result from of one of the receiver is shown in Fig. 5. As can be seen in Fig. 5, if the medium delay setting is swept from 0 to 63, then the changes of delay value is about 0 to 2.98 ns ($\approx 63 \times 45$ ps).

C. Near Field Beamforming

After array calibration, the point source radiator is located at different points at depth of 50 cm in front of the array. Due to the limited azimuth angle of Vivaldi antennas and free-space path loss, the receiving array scans from -70 cm to 70 cm at depth of 50 cm.

To scan each point on the corss-range line, the receiving elements are delayed accordingly to have constructive interference in that point. To illustrate the array performance, the point source is located on the array center line and 30 cm off the center line, and the normalized measured electromagnetic (EM) radiations are shown in Fig. 6. As shown in Fig. 6 the measured EM radiation is strong in the area with the darkest and lightest colors. As can be seen in the top and the bottom images of Fig. 6, the measured EM radiation is stronger around center line and 30 cm off the array center line, respectively. As presented in Fig. 6, the measured electromagnetic radiation in near-field is spherical, but it is plane-wave in the far-field.

The measured energy patterns of the array when the point source is located on and 30 cm off the array center line are shown in Fig. 7. As shown in Fig. 7(a) the beamwidth at half of the mainlobe is about 17 cm. The measured ISL which is the ratio of the largest sidlobe level to mainlobe level ratio is about: $10 \log_{10} \frac{0.3}{1} = -5.3$ dB.

V. CONCLUSION

An IR-UWB beamforming antenna array is presented. The array is implemented with 5 impulse radio radar transceiver

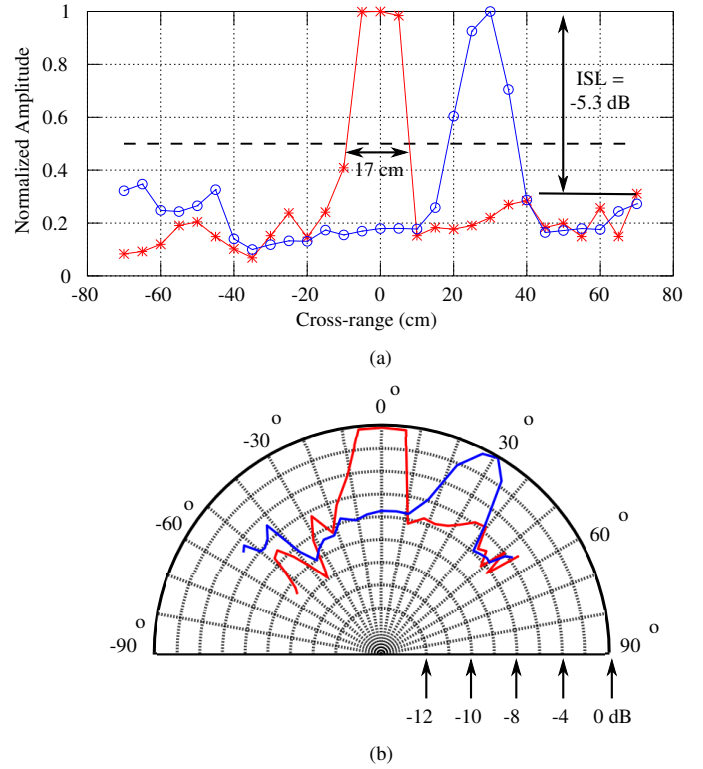


Fig. 7. (a) Measured energy pattern in rectangular coordinate. (b) Measured energy pattern in polar coordinate.

and 10 cm antenna spacing. The received signal can be recovered with high resolution using a high speed sampler with sampling rate of 10 GHz. The electronic steering is based on the delay and sum beamforming technique. For accurate delay tuning and signal alignment a novel calibration circuit and procedure is used. The measured energy pattern has a beamwidth of 17 cm. The proposed antenna array may be suitable for medical imaging and biomedical application.

REFERENCES

- [1] H. Hashemi and H. Krishnaswamy, "Challenges and opportunities in ultra-wideband antenna-array transceivers for imaging," in *Proc. IEEE ICUWB*, Sep. 2009, pp. 586–591.
- [2] T.-S. Chu *et al.*, "An integrated ultra-wideband timed array receiver in 0.13 μ m cmos using a path-sharing true time delay architecture," *IEEE J. Solid-State Circuits*, vol. 42, no. 12, pp. 2834–2850, Dec. 2007.
- [3] F. Bautista *et al.*, "UWB beamforming architecture for rtls applications using digital phase-shifters," in *Proc. IEEE ISCAS*, May 2011, pp. 1540–1543.
- [4] H. A. Hjortland and T. S. Lande, "CTBV integrated impulse radio design for biomedical applications," *IEEE Transactions on Biomedical Circuits and Systems*, vol. 3, no. 2, pp. 13–23, Apr. 2009.
- [5] D. H. Johnson and D. E. Dudgeon, "Array signal processing- concepts and techniques," in *Prentice Hall*, 1993.
- [6] X. Zhuge and A. G. Yarovoy, "A sparse aperture mimo-sar-based uwb imaging system for concealed weapon detection," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, no. 1, pp. 509–518, Jan. 2011.
- [7] K. K. Lee *et al.*, "A 5.2 pJ/pulse impulse radio pulse generator in 90 nm cmos," in *Proc. IEEE ISCAS*, 2011, pp. 1299–1302.
- [8] T. A. Vu *et al.*, "Continuous-time cmos quantizer for ultra-wideband applications," in *Proc. IEEE ISCAS*, Jun. 2010, pp. 3757–3760.
- [9] M. Zarre Dooghabadi *et al.*, "High precision calibrated digital delay element," *Electronics Letters*, vol. 47, no. 9, pp. 564–565, Apr. 2011.

Paper 17 A linear IR-UWB MIMO radar array

M. Z. Dooghabadi, H. A. Hjortland, and T. S. B. Lande. “A linear IR-UWB MIMO radar array”. In: *Ultra-Wideband (ICUWB), 2013 IEEE International Conference on*, pp. 13–19, Sydney, NSW, Sep. 2013. doi:10.1109/ICUWB.2013.6663814. © 2013 IEEE. Reprinted with permission.

My Contributions to This Paper

This paper presents a radar system with beamforming. I came up with much of the design for the radar system which is being used here, and I also came up with most of the design for the calibration circuit which enables the precise beamsteering. I took part in discussions about the design of the beamforming experiment, my SPI and microcontroller solutions were used, and I also helped review the paper.

A Linear IR-UWB MIMO Radar Array

Malihe Zarre Dooghabadi, Håkon A. Hjortland, and Tor Sverre “Bassen” Lande

Department of Informatics, University of Oslo, Norway, Email: malihezd@ifi.uio.no

Abstract—A linear impulse radio ultra wideband MIMO radar array is presented. The IR-UWB imaging array is true time delay based. The imaging system is implemented using an impulse radio radar transceiver which is realized in 90 nm CMOS technology. The transceiver has programmable coarse, medium and fine delay lines which facilitate an electronic beamsteering with high resolution. A calibration circuit and scheme for precise delay adjustment of antenna elements are presented. Tapered slot Vivaldi antennas are used in both of the transmit and receive arrays. The MIMO radar array is modeled in MATLAB. Using the accurate calibration procedure, measurement results agree very well with the simulation results. A clutter map removal method is used for reducing the antenna elements cross-talk effect and the static variations to improve the image quality. The effects of different array parameters such as inter element spacing and number of transmit and receive antenna elements on the array pattern in terms of resolution and signal to noise ratio are studied to optimize and enhance the MIMO radar array performance.

I. INTRODUCTION

The low cost and low power impulse-radio ultra wideband (IR-UWB) radar transceiver array with multiple transmit and receive antenna elements has potential applications in medical imaging and surveillance systems. The short duration and wideband Gaussian waveforms provide high cross-range resolution. In UWB imaging arrays, due to the short duration of impulses there will be less interferences outside the focusing area, so the spacial sampling can be larger than the half of the wavelength without high level grating lobes [1], [2]. Therefore, the UWB imaging arrays are less dense and low cost compared to the conventional narrowband phased arrays.

The IR-UWB imaging arrays are true time delay (TTD)-based [3]. Several TTD-based beamforming systems have been presented in [4]–[7]. The presented UWB beamforming systems in [4]–[7] are either transmitting or receiving arrays. In [4], [5] on-chip coplanar strip line and tapped delay trombone line are used for delay adjustment, but this architecture requires a large number of on-chip inductors, transmission lines, and complex trombone tap circuit structures, and the reported true time delay and phase resolution are limited to 17.5 ps and 10°, respectively. In [6] a 4-channel beamforming transmitter is presented with fine delay resolution of 10 ps and phase resolution of 1°. The presented UWB four-element transmitting array in [7] is implemented with linear delay lines, micro-strip UWB power dividers and broadband transistors, and a wideband digital sampling oscilloscope is used as a receiver.

In this paper a Multiple-Input Multiple-Output (MIMO) array using several IR-UWB radar transceivers is presented. The transceiver circuit is based on the Continuous-Time Binary-Value (CTBV) signal processing technique [8], [9]

which facilitates a simple and low cost design using digital circuits. In the CTBV technique there is no running clock signal and timing is based on the inherent, process-dependent gate delays [8]. In the CTBV signal processing domain, the signal amplitude is a binary value of either 1 or 0, and it is continuous in time. The radar transceiver has a high speed sampler which is not clocked and uses continuous-time signal processing that provides a sampling rate of 10 GHz which recovers the backscattered signal with high resolution [10].

The digital programmable delay elements are used as the true time delay elements with coarse, medium and fine resolutions of 900 ps, 45 ps, 10 ps, respectively.

An important design parameter in TTD-based imaging arrays is accurate delay adjustment to have electronic beamsteering with high resolution. In order to compensate all the deviations and device mismatches in the system, the delay elements must be calibrated. The radar transceiver has a calibration circuit with high precision for accurate delay adjustment [11]. Also, a calibration procedure is used to compensate the mismatches between chips, PCBs, cable, and antennas for aligning all the transmit and receive antenna elements.

The linear MIMO radar array is implemented with different numbers of transmit/receive antenna element and various inter element spacing to study the effects of these design parameters on the beampattern and array imaging performance.

The rest of the paper is organized as follows. In section II the delay and sum beamforming technique and design considerations in imaging arrays are investigated. The IR-UWB MIMO radar array implementation and imaging are explained in sections III and IV. Simulation and experimental results are presented in section V, and section VI concludes the paper.

II. DELAY AND SUM BEAMFORMING AND ANTENNA ARRAY CHARACTERIZATION

The presented IR-UWB MIMO array is based on the delay and sum beamforming technique. In delay and sum beamforming the transmitting/receiving elements are delayed to have constructive interference at the desired point or direction. The output of a linear array with n antenna elements is [12]:

$$y(t) = \sum_{i=0}^{n-1} w_i s_i(t - \Delta_i) \quad (1)$$

where $s_i(t)$ and w_i are impulse signal and element weighting coefficient, respectively. The Δ_i is the steering delay. The

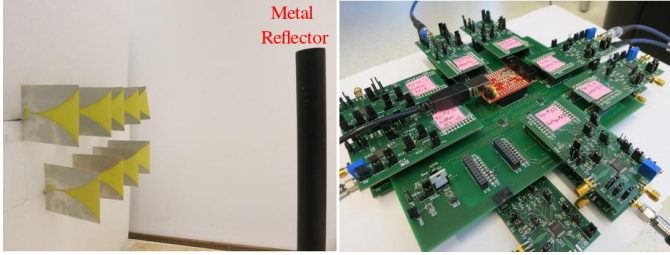


Fig. 1. IR-UWB MIMO radar array setup and beamforming system.

energy pattern of an array is defined as [13], [14]:

$$P_{\text{energy}}(\phi) = \int_0^\infty |y(t, \phi)|^2 dt \quad (2)$$

To measure the array ability to reject the unwanted signal the sidelobe level is defined as [14]:

$$SLL = \frac{\text{maximum value of largest sidelobe}}{\text{maximum level of mainlobe}} \quad (3)$$

The cross-range resolution depends on the array size and the center frequency [1]:

$$\delta = \frac{\lambda_c R}{L} \quad (4)$$

where R is the distance between array and target, λ_c is wavelength at the center frequency, and L is the array size. So, an array with the larger aperture size and higher center frequency results in a beampattern with narrower beamwidth and better resolution. The cross-range direction is perpendicular to the array center line.

In conventional narrowband beamforming to avoid the spatial aliasing and the grating lobes the antenna element spacing must not be more than the half of the wavelength. But, in UWB beamforming due to the short pulses there will be less interferences outside the focusing location [1], [2]. To reduce the cost and power consumption, the UWB arrays can have few antenna elements with inter element spacing larger than the half of the wavelength without the grating lobe. However, as the array gets sparser the sidelobes raise more.

III. IR-UWB MIMO RADAR ARRAY IMPLEMENTATION

A. MIMO Radar Array Architecture

The implemented linear MIMO antenna array consists of linear transmitting and receiving arrays. Tapered slot Vivaldi antenna are used by both of the transmitters and receivers. Vivaldi antenna are wideband and have directive patterns which facilitates to focus energy along one direction instead of all directions [15], [16]. The Vivaldi antennas which are used in the MIMO radar array have a gain of $\approx 10\text{dBi}$ at the center frequency of about 2.5 GHz, and dimension of $10\text{ cm} \times 15\text{ cm}$ [15]. The Vivaldi antennas which are used in the measurement setup have wide azimuth angle and narrow elevation angle. Due to their wide azimuth angle, if the transmit and receive antennas are placed beside each other on the azimuth direction, there will be a strong cross-coupling between them. In order to reduce the cross-talk between the

transmit and receive antenna elements, the transmit and receive arrays are placed above each other as shown in Fig. 1.

The IR-UWB beamforming system as presented in Fig. 1 is implemented with several IR-UWB radar transceiver modules. In radar transceiver chip there is an SPI interface which is used for register settings and to generate the controlling signals. The micro-controller which is programmed through USB interface is used for signal generation, SPI programming, data collection and readout.

B. IR-UWB Radar System

Fig. 2 shows the block diagram of the IR-UWB radar transceiver which has been presented in [10]. The presented IR-UWB radar is used for UWB beamforming. The radar transceiver is fabricated in 90 nm TSMC technology and it is based on the Continuous-Time Binary-Value (CTBV) signal processing [8], [9]. Using the CTBV signal processing technique, the IR-UWB transceiver system except the pulse generator and the quantizer is implemented with simple digital circuits.

The transmitter has two main circuits of programmable delay element and a pulse generator which generates high order Gaussian pulses [17]. In radar transceiver the programmable delay elements are used for ranging in both of the transmitter and receiver circuits. The programmable delay element of the transmitter circuit consists of fine and medium delay lines. The fine and medium delay lines have 65 delay settings with tuning steps of 10 ps and 64 delay settings with tuning steps of 45 ps, receptively. By tuning the programmable delay element the pulse transmission of the different transmitters may be coordinated for beamforming. The programmable delay element of the receiver has medium and coarse delay lines. The medium and coarse delay lines have 64 delay settings with step size of 45 ps and 900 ps, receptively. By programming the delay lines the receiver can listen to different ranges.

The receiver circuit consists of a 1-bit quantizer, a programmable delay element, a digital sampler, a digital integrator, and a readout circuit. The 1-bit quantizer with tunable threshold setting quantizes the backscattered signal to binary value of either 1 or 0. The digital output of the quantizer which is a CTBV coded signal is sampled by a continuous-time sampler. The sampler has 64 D flip-flops as sampling elements which tap a chain of delay elements in steps of 100 ps that provides a sampling rate of 10 GHz. Using the presented sampler with sampling frame of about 6 ns the backscattered signal can be recovered in the depth dimension.

The digital integrator is implemented using counters which count the number of ones. By sweeping the threshold voltage combined with integration the received signal can be recovered with high resolution. The amount of integrations or processing gain is controlled by the *Trigger* signal.

In transmit beamforming by adjusting the delay values, the pulse transmission may be set up for constructive interferences at some focal point, or in receive beamforming the receivers can be aligned accordingly to listen to the desired direction.

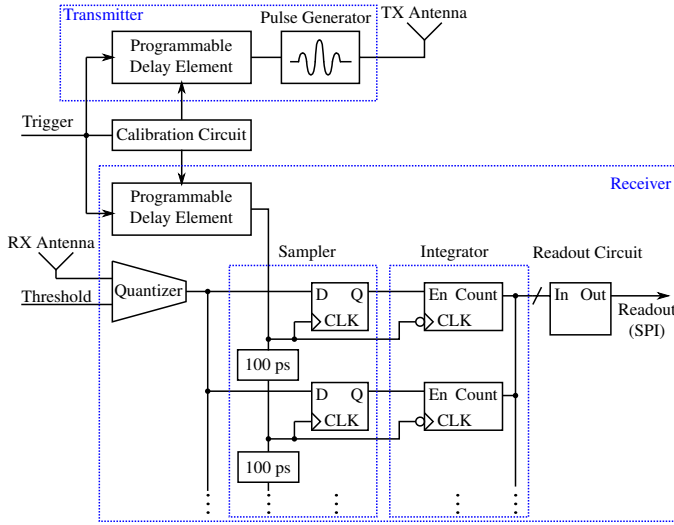


Fig. 2. IR-UWB radar transceiver.

So, the delay adjustment with high precision is an important design parameter in delay and sum beamforming, and hence the delay elements must be calibrated. A calibration circuit and procedure which was presented in [11] is used for precise delay adjustment and to compensate all the mismatches and deviations.

The calibration circuit block diagram is shown in Fig. 3. In calibration mode after enabling the calibration loop which basically is a gated ring oscillator, it starts to oscillate. Then, as long as the reference pulse is high, the loop counter counts the oscillation cycles. The loop signal period is the reference pulse width divided by the loop count values. The loop signal period counts for all the delay elements inside the calibration loop. The time diagram in Fig. 3. presents the circuit functionality.

For electronic beamsteering, if the transmit or receive antenna elements need to be delayed +100 ps or -100 ps, and the loop signal period is for example 10 ns, then the programmable delay element is swept and the calibration procedure is repeated until the loop signal period is 10.1 ns or 9.9 ns. So, this way the delay elements can be adjusted with precision of few pico seconds RMS error [11].

IV. IR-UWB MIMO RADAR ARRAY IMAGING

In delay and sum beamforming all the transmitting and receiving elements must be aligned and calibrated. By calibration, the transistor mismatches and process variations inside the transceiver circuits and also all the mismatches on PCBs, cables, and antennas can be compensated.

Calibration procedure is done in two steps, first all the transmitting and receiving elements are aligned according to a reference point. Then, for electronic beamsteering and scanning the cross-range line at the desired depth, the required delay values for each of the antenna elements are calculated. Using the calibration circuit and calibration scheme which are explained in section III, the correct delay settings for the fine, medium, coarse delay lines can be found.

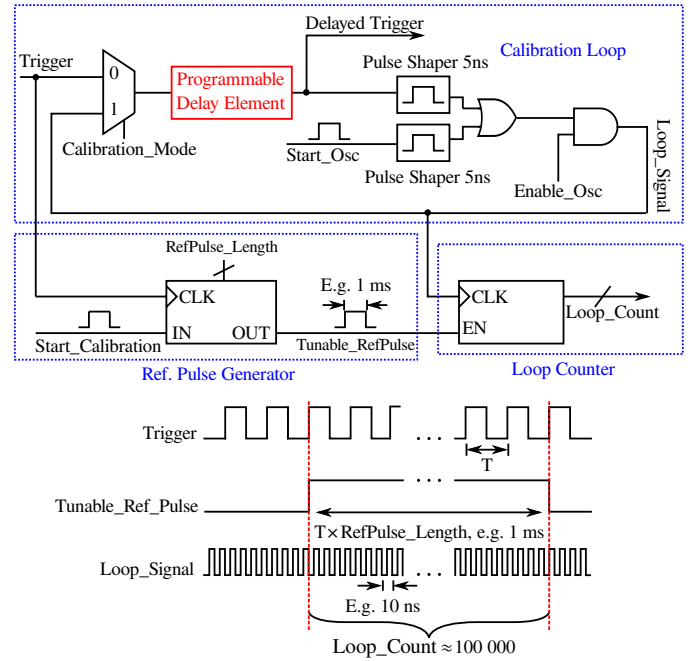


Fig. 3. Calibration circuit and calibration timing diagram.

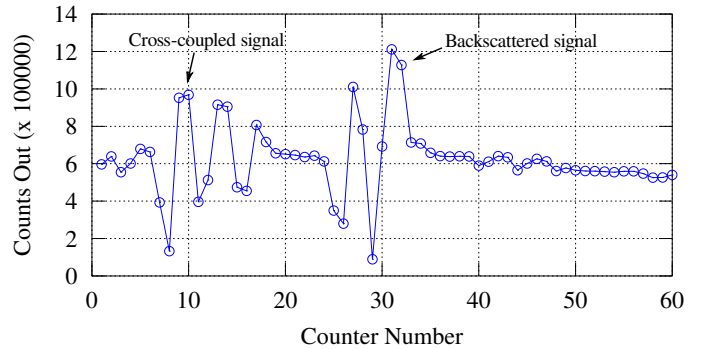


Fig. 4. Measured readout of one of the receiver after calibration.

A. Array Calibration

To calibrate and align the antenna elements a metal reflector is placed in front of and on the center line of the MIMO radar array, as a reference or calibration point. This setup is shown in Fig. 1. First of all, the receiving elements must be tuned to listen to the depth in which the target is located. To align the receivers, a transmitter which is in the center of the transmitting array emits, and then the programmable delay element of each of the receivers is swept until the received signal is fitted in the middle of the sampling frame.

With this procedure all the receiving elements are aligned. Fig. 4 shows one of the receiver readout after calibration, as can be seen in this figure the backscattered signal is fitted in the middle of the sampling frame. The signal which appears at the start of the sampling frame is due to the cross-coupling between transmit and receive antenna elements.

After all the receivers are calibrated and aligned, the transmitting elements must be calibrated too. In transmitting

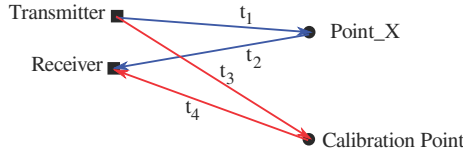


Fig. 5. The time difference is: $\Delta T = (t_1 + t_2) - (t_3 + t_4)$.

calibration procedure, just one calibrated receiving element is enabled. Then each of the transmitters emits and its programmable delay element is swept, until the backscattered signal appears in the middle of the sampling frame of the enabled receiver. In this way the transmitting array and the enabled receiving element are calibrated accordingly. If the calibrated transmitters fire, the transmitted signals will interfere constructively at the calibration point, and the backscattered signal will appear in the middle of that receiver sampling frame.

The transmitting calibration scheme is repeated with other calibrated receivers to make a calibration table for the transmitting array. The calibration table is used to make a scan table for scanning the cross-range line at the desired depth from array.

B. Electronic Beamsteering

Using transmitting/receiving array calibration procedures, all the chips, PCBs, cables, and antennas mismatches will be corrected and the antenna elements are aligned and ready for electronic beamsteering.

For imaging and electronic beamsteering the target is placed at the desired depth and in front of the MIMO radar array, then the transmitters sweep the cross-range line and the receivers receive the backscattered signal, so where the target is located there will be a strong reflection and in other points we expect no reflections.

For scanning all the points on the cross-range line and transmit sweeping, a scan table is made using the calibration table. The time of flight from transmitter to a point on the cross-range line and then from that point to the receiver is calculated. Also, the time of flight from that transmitter to the reference or calibration point, and then from calibration point to that receiver is calculated. As illustrated in Fig. 5, by subtracting these time of flights the required delay for beamsteering is found. According to the calculated delay difference the delay value of the transmitter is decreased or increased, and its programmable delay element is adjusted precisely using the calibration circuit. A scan table with required delay settings for electronic beamsteering all transmitters can be created using this procedure.

V. SIMULATION AND MEASUREMENT RESULTS

The linear MIMO radar array is implemented and calibrated as explained in previous sections. To investigate the effects of different design parameters such as antenna spacing, and number of antenna elements on the array pattern, the MIMO radar array is configured with different settings. Also, the beamforming system is modeled in MATLAB. Simulation and measurement results are compared.

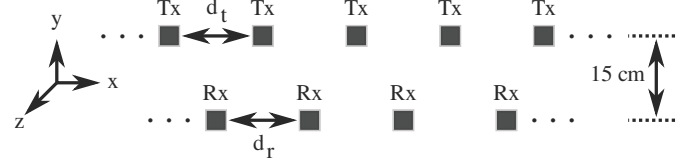


Fig. 6. MIMO radar array measurement/simulation setup. Different number of transmitters and receivers are used for different test setups.

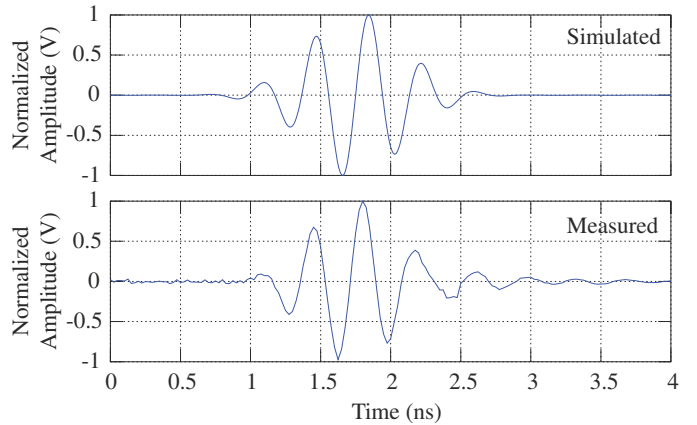


Fig. 7. Simulated and measured Gaussian pulse.

A. Measurement Setup

The measurement setup of the IR-UWB MIMO radar array is shown in Fig. 6. As explained in previous sections, the transmit and receive antenna elements are placed above each other to reduce the cross-coupling effect. In the receive array the inter element spacing is set to 10 cm which is about one wavelength. To investigate the effect of antenna spacing on array pattern the inter element spacing of the transmit array is set 10 cm and 20 cm.

B. Simulation Setup

In simulation the Gaussian pulse is modeled as:

$$G(t) = e^{\frac{-t^2}{2u_B}} \times \sin(2\pi f_c t + \phi) \quad (5)$$

where $u_B = \frac{1}{\pi f_B}$, $f_B = \frac{BandWidth}{2}$, f_c is the center frequency, and ϕ is the phase shift. The variables in equation 5 are set according to the measured center frequency, measured pulse duration, and measured amplitude. Both of the simulated and measured Gaussian pulses are shown in Fig. 7. The LeCroy oscilloscope with sampling rate of 80 Gsamples/s is used to measure the Gaussian pulse. The array output is modeled based on equation 1. The arrays geometry, number of antenna elements, and the antenna spacing in the modeled MIMO array is the same as the measurement MIMO radar array setup.

C. MIMO Radar Array Imaging

The MIMO radar array is configured as illustrated in Fig. 6. The transmit array has 3 antenna elements with inter element

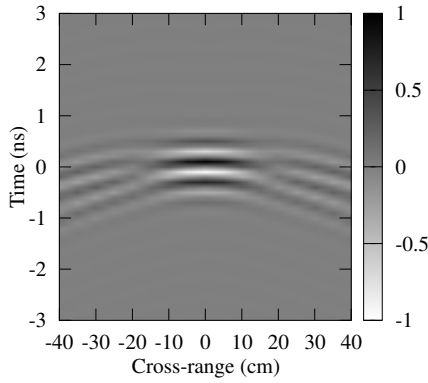


Fig. 8. Simulated output of the MIMO radar array based on equation 1. The metal reflector is at the center.

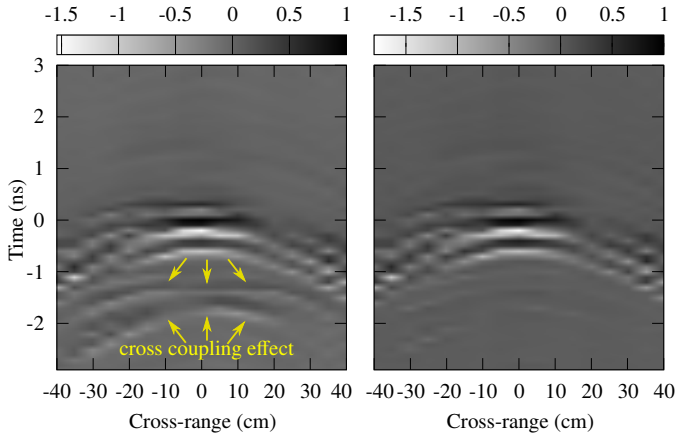


Fig. 9. Measured output of the MIMO radar array based on equation 1. The metal reflector is at the center.

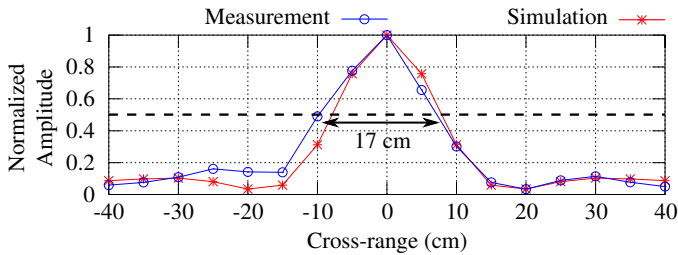


Fig. 10. Measured/simulated energy pattern. The metal reflector is located on the center line of the array.

spacing of 10 cm, and the receive array has 4 antenna elements with antenna spacing of 10 cm. All the transmitters and receivers are calibrated and aligned as explained in IV.

In order to evaluate the imaging system functionality a metal reflector is placed in front of the MIMO radar array at the depth of 50 cm from transmit/receive arrays. Since the Vivaldi antennas which are used in the measurement setup are not ideal and have limited azimuth angles, the cross-range line at the depth of 50 cm is swept from -40 cm to 40 cm.

Fig. 8 and Fig. 9 show the simulated and measured outputs of the MIMO radar array with above settings and when the metal reflector is placed on the center of the MIMO radar

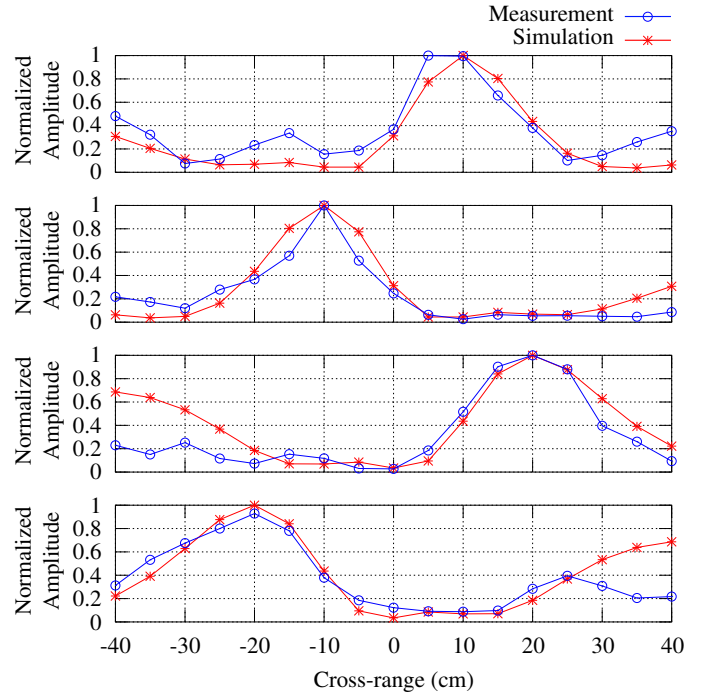


Fig. 11. Measured/simulated energy patterns when the metal reflector is at different positions (± 10 cm and ± 20 cm of the array center line).

array. The backscattered signal is strong in the the area with darkest and lightest colors which indicates that the target is located there.

As mentioned in previous sections, there are cross-talks among transmit and receive antennas. This cross-coupling effect is shown in the left image of Fig. 9. In order to get rid of the cross-coupling effect, a clutter map removal method is used. First the target is removed and the MIMO radar array sweeps the cross-range line to make an empty-room image. After that the target is placed in front of the array and the MIMO radar array scans the cross-range line, then the empty-room image is subtracted from the obtained image to remove the cross-coupling effect. The right image in Fig. 9 is after empty-room image subtraction, and as can be seen the cross-coupling effect is reduced and the image quality is improved. The energy pattern is presented in Fig 10. As can be seen in the energy pattern the measured beamwidth at half of the mainlobe height or half-power beamwidth (HPBW) [14] is about 17 cm.

As presented in figures 8, 9 and 10, simulation and measurement results are similar, and it is because of the calibration procedure and high precision calibrated delay elements, which are necessary in delay and sum beamforming.

Fig. 11 presents the energy patterns of the array, when the metal reflector is placed at different positions. As can be seen in these graphs when the target is off the array center line, the beamwidth of the main lobe and the level of side lobes increase. When the target is off center the time of flight is increased giving an attenuated signal, and hence the sidelobe level (SLL) degrades.

D. Array Pattern Characterization

The array parameters such as inter element spacing, number of transmitting and receiving antenna elements can affect the array pattern. So, the array performance in terms of resolution, sidelobe level, and signal to noise ratio can be optimized based on these design parameters.

1) *Antenna Spacing*: As stated in equation 4 by increasing the aperture size the mainlobe in the array pattern gets narrower, and hence the cross-range resolution is improved. The energy patterns of two MIMO radar arrays with the same number of transmitting and receiving elements but different antenna spacings of 10 cm and 20 cm are shown in Fig. 12(a). As can be seen in these graphs the array with inter element spacing of 20 cm has narrower mainlobe and better resolution than the array with antenna spacing of 10 cm. However, as the arrays becomes sparser, the level of sidelobes increases.

2) *Number of Transmitting Antenna Elements*: The energy patterns of the MIMO arrays with the same size but different number of transmitting elements of 5 and 3 are shown in Fig. 12(b). As can be seen by thinning and removing the transmitting elements within the same aperture size, the array becomes sparser. Due to thinning the level of sidelobes increases, but the mainlobe does not change and the beamwidth is marginally improved.

3) *Number of Receiving Antenna Elements*: As mentioned in section IV, the transmit array scans and sweeps the cross-range line and the receive array receives the backscattered signal. Therefore, more receiving elements improves the signal strength and the signal to noise ratio. The array patterns of two MIMO radar arrays with 4 and 2 receiving elements are shown in Fig. 12(c). As presented in Fig. 12(c) the MIMO radar array with more receiving elements has a stronger mainlobe, and signal to noise ratio is improved.

E. Multiple Objects Imaging

According to the presented measurement results in section V-D, the MIMO radar array with 5 transmitting and 4 receiving antenna elements and 10 cm inter element spacing has the best performance in terms of resolution and sidelobe level. As shown in Fig. 12(b), the array has a beamwidth of 10 cm, so it can distinguish multiple objects when the spacing between them is more than 10 cm.

Two metal objects with 20 cm spacing are placed in front of the MIMO radar array. The array output and the energy pattern are shown in Fig. 13 and Fig. 14, respectively. As illustrated in Fig. 14, the two mainlobe peaks are above the HPBW line and the null between the two mainlobe is under the HPBW line, so this indicates that the array distinguishes two separate objects.

As can be seen in Fig. 14 the sidelobe level raises compared to the Fig. 11 where only one object is located 10 cm off the array center line. This is due to the fact that with more objects the interference level increases which results to the high level sidelobes.

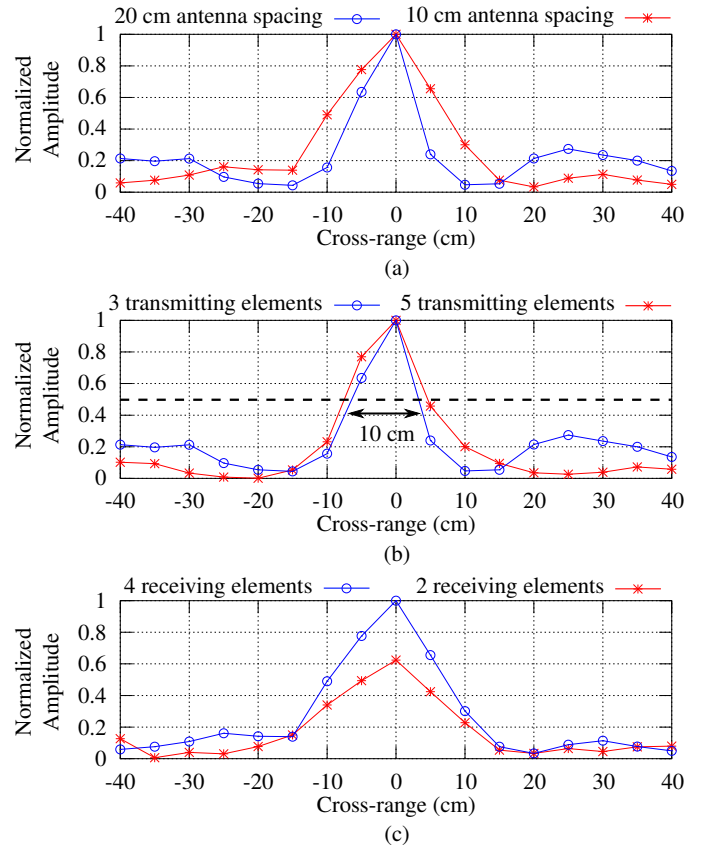


Fig. 12. Measured energy patterns for: (a) different aperture sizes, (b) different number of transmitting elements, (c) different number of receiving elements.

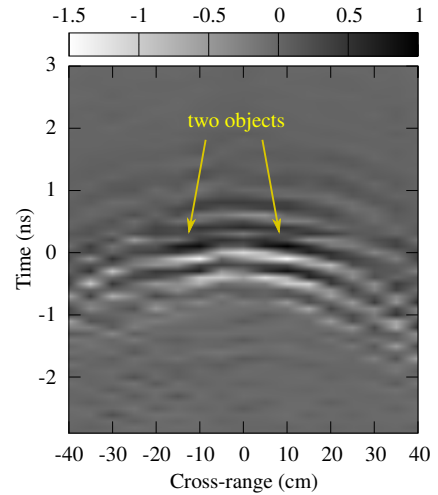


Fig. 13. Measured output of the MIMO radar array based on equation 1. Two metal objects are located in front of the array.

VI. CONCLUSION

An impulse radio ultra wideband MIMO radar array for close range imaging is presented. The MIMO radar array is based on the delay and sum beamforming. The Imaging system is implemented using several single chip radar transceivers

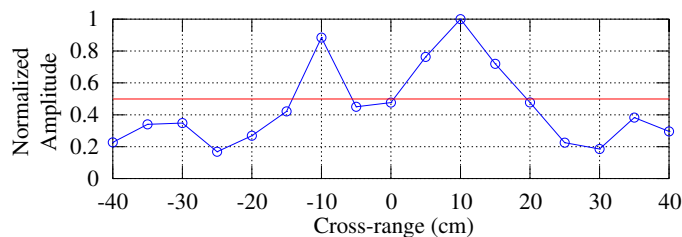


Fig. 14. Measured energy pattern when two metal objects are located in front of the array.

based on Continuous-Time Binary-Value signal (CTBV) processing.

Combining the CTBV signal processing technique which facilitates a sampling rate of 10 GHz with calibration circuit and calibration scheme which provides delay adjustment with precision of few pico seconds, accurate electronic beamsteering is feasible. By calibration most of the static deviations in the beamforming system can be compensated.

Due to the short impulses, even in close range a clutter map removal method can be used to reduce the cross-coupling between the transmit and receive antennas and to improve the image quality.

The MIMO radar array is configured with different number of transmit and receive antenna elements and various inter element spacings to investigate the array focusing performance. In UWB array the antenna spacing can be larger than the half of the wavelength, however the reduced spacial sampling results in higher sidelobe levels. The presented MIMO radar array may have sufficient scanning capability for close range imaging.

REFERENCES

- [1] X. Zhuge and A. G. Yarovoy, "A sparse aperture mimo-sar-based uwb imaging system for concealed weapon detection," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, no. 1, pp. 509–518, Jan. 2011.
- [2] S. Ries and T. Kaiser, "Ultra wideband impulse beamforming: it is a different world," *Signal Processing*, vol. 86, no. 9, pp. 2198–2207, Sep. 2006.
- [3] H. Hashemi, T.-S. Chu, and J. Roderick, "Integrated true-time-delay-based ultra-wideband array processing," *IEEE Communications Magazine*, vol. 46, no. 9, pp. 162–172, Sep. 2008.
- [4] T.-S. Chu, J. Roderick, and H. Hashemi, "An integrated ultra-wideband timed array receiver in 0.13 μm cmos using a path-sharing true time delay architecture," *IEEE J. Solid-State Circuits*, vol. 42, no. 12, pp. 2834–2850, Dec. 2007.
- [5] Z. Safarian, T.-S. Chu, and H. Hashemi, "A 0.13 μm cmos 4-channel uwb timed array transmitter chipset with sub-200ps switches and all-digital timing circuitry," in *IEEE Radio Frequency Integrated Circuits Symposium*, Apr. 2008, pp. 601–604.
- [6] L. Wang, Y. X. Guo, Y. Lian, and C. H. Heng, "3-to-5ghz 4-channel uwb beamforming transmitter with 1 phase resolution through calibrated vernier delay line in 0.13 μm cmos," in *IEEE International Solid-State Circuits Conference Digest of Technical Papers (ISSCC)*, Feb. 2012, pp. 444–446.
- [7] C.-H. Liao, P. Hsu, and D.-C. Chang, "Energy patterns of uwb antenna arrays with scan capability," *IEEE Transactions on Antennas and Propagation*, vol. 59, no. 4, pp. 1140–1147, Apr. 2011.
- [8] H. A. Hjortland and T. S. Lande, "CTBV integrated impulse radio design for biomedical applications," *IEEE Transactions on Biomedical Circuits and Systems*, vol. 3, no. 2, pp. 13–23, Apr. 2009.
- [9] Y. W. Li, K. L. Shepard, and Y. P. Tsividis, "Continuous-time digital signal processors," in *IEEE International Symposium on Asynchronous Circuits and Systems*, Mar. 2005, pp. 138–143.
- [10] M. Z. Dooghabadi, H. A. Hjortland, and T. S. Lande, "An ultra-wideband receiving antenna array," in To be published at: *Proc. IEEE ISCAS*, 2013.
- [11] M. Z. Dooghabadi, H. A. Hjortland, and T. S. Lande, "High precision calibrated digital delay element," *Electronics Letters*, vol. 47, no. 9, pp. 564–565, Apr. 2011.
- [12] D. H. Johnson and D. E. Dudgeon, "Array signal processing- concepts and techniques," in *Prentice Hall*, 1993.
- [13] X. H. Wu, A. A. Kishk, and Z. N. Chen, "A linear antenna array for uwb applications," in *IEEE International Symposium on Antennas and Propagation Society*, vol. 1A, Jul. 2005, pp. 594–597.
- [14] M. Hussain and A. S. Al-Zayed, "Aperture-sparsity analysis of ultrawideband two-dimensional focused array," *IEEE Trans. Antennas Propag.*, vol. 56, no. 7, pp. 1908–1918, Jul. 2008.
- [15] T. A. Vu, M. Z. Dooghabadi, S. Sudalaiyandi, H. A. Hjortland, O. Nass, T. S. Lande, and S. E. Hamran, "Uwb vivaldi antenna for impulse radio beamforming," in *NORCHIP Conference*, November 16th to 17th 2009, Trondheim, Norway, 2009.
- [16] P. J. Gibson, "The vivaldi aerial," in *Microwave Conference, 1979. 9th European*, Sep. 1979, pp. 101–105.
- [17] K. K. Lee, M. Z. Dooghabadi, H. A. Hjortland, O. Nass, and T. S. Lande, "A 5.2 pj/pulse impulse radio pulse generator in 90 nm cmos," in *Proc. IEEE ISCAS*, 2011, pp. 1299–1302.

A.4 Bitstream Processing

Paper 18 Power efficient cross-correlation beat detection in electrocardiogram analysis using bitstreams

O. E. Liseth, H. A. Hjortland, and T. S. Lande. “Power efficient cross-correlation beat detection in electrocardiogram analysis using bitstreams”. In: *Biomedical Circuits and Systems Conference, 2009. BioCAS 2009. IEEE*, pp. 237–240, Beijing, Nov. 2009. doi : 10.1109/BIOCAS.2009.5372041. © 2009 IEEE. Reprinted with permission.

My Contributions to This Paper

This is a conference paper which was later invited for the special journal issue from the conference. As co-supervisor of these MS students I worked with them during development and implementation of a running cross correlator. Somewhat involved in the actual writing as well.

Power Efficient Cross-Correlation Beat Detection in Electrocardiogram Analysis Using Bitstreams

Olav E. Liseth, Håkon A. Hjortland and Tor Sverre “Bassen” Lande

Dept. of Informatics, University of Oslo, Norway

Email: olaveli@ifi.uio.no, haakoh@ifi.uio.no, bassen@ifi.uio.no

Abstract—In this paper we present a novel cross-correlator chip suitable for “smart” ECG electrodes. Sophisticated QRS-detection is feasible exploring multi-component based cross-correlation and by exploiting the simplicity offered by bitstream processing. The chip is evaluated using real ECG signals as an example application. Power-efficient running cross-correlation is obtained with novel asynchronous circuit solutions consuming only approx. 5 mW. The implemented generic cross-correlator may be explored for other pattern-matching applications in wireless sensor networks or other low-power applications.

I. INTRODUCTION

The notion of small, portable, battery-operated systems often organized in a wireless sensor network (WSN) has initiated significant research activity. Usage of WSNs in body sensor networks (BSN) [1] is interesting and may facilitate elegant and compact solutions for vital sign monitoring. In BSN applications, size is limited and low-power operation is mandatory for battery operation. The benefit of adopting specialized silicon systems for minimal size and power consumption in BSN applications is evident [2].

An interesting application of BSNs is personalized health monitoring. Small sensor nodes may be attached to the body and even located inside the body as implants. Organized as WSNs, health information is processed, conditioned and wirelessly interacting with surrounding networks.

An important vital sign indicator is the heartbeat or the more accurate analysis of the electrical heart activity known as ECG analysis. Measuring of heartbeats of the simplest form may be done by sensing an artery’s pulsation at the body surface and is merely usable for heart rate measurements. The elaborate ECG-analysis requires significant bedside equipment and typically 12 electrodes attached to the body. These procedures are often carried out in hospitals.

However, there is an intermediate ECG analysis often used for long-term observations known as Holter monitors. These systems tend to have fewer electrodes, normally 2–3 attached to the torso. The Holter monitor is worn during normal activity and should not interfere with normal lifestyle. These ECG-light monitors do catch serious cardiac arrhythmia like Bradycardia, Tachycardia, Sinus arrest, Ventricular tachycardia, Supraventricular tachycardia and similar. Older systems left hours of ECG signals to be analyzed by cardiologists, but over the year a number of methods have been proposed for computerized ECG analysis [3].

In this paper we are proposing a novel single-chip cross-correlator which can be used to perform ECG analyses. A pro-

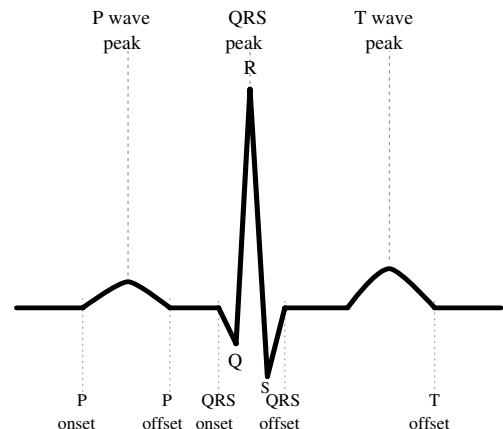


Fig. 1. Schematic representation of a normal ECG waveform including P wave, QRS complex and T wave beat markers.

grammable “smart” ECG electrode with embedded heartbeat detection or even more elaborate ECG analyses is feasible. Combined with body area network technology, a wireless ECG system for long-term usage may be constructed.

II. HEARTBEAT DETECTION

Beat detection involves identifying all cardiac cycles in ECG recordings and locating each identifiable waveform component within a cycle, the P wave peak, the QRS peak and the T wave peak, including the associated on- and offset timing. The electrical activity from one cycle of cardiac contraction and relaxation is schematically drawn in Fig. 1. The accuracy of the beat detection has significant impact on the overall classification performance and various computerized methods have been developed over the years trading off computational efficiency with detection quality.

A. Multi-component Based Heartbeat Detection

Although cross-correlation of the ECG-signal has been explored for template-matching ECG-signals with anomaly templates, this method was abandoned due to both poor performance and high computational demands. Recently an improvement of the traditional cross-correlation method was proposed in [4], where each of the identifiable waveform components within a cycle are searched for in isolation. Three templates were used, one for the P wave, one for the QRS complex and one for the T wave. The first step is to locate the QRS complex by cross-correlating the QRS template with

the ECG signal. This method is repeated with the P and the T wave templates. In each case, the highest value of the cross-correlation result within the given correlation interval is compared to a predefined threshold. The threshold value is established during a pre-learning phase and can be adjusted during operation. Although the component-based heart-beat detection is functionally promising, the computational complexity of cross-correlation is still demanding. In WSN applications, power-efficient implementations are required.

III. BITSTREAM CROSS-CORRELATOR

A discrete estimate of the cross-correlation between two sequences is found by the equation:

$$r(t) = \sum_{k=0}^{n-1} y(k)x(t+k)$$

The sequence $x(t)$ is cross-correlated with a template sequence, $y(t)$, by multiplying each element of the two sequences over a window of length n and summing the result. The result, $r(t)$, is a good estimate of the cross-correlation between the two sequences. For every new sample of the incoming signal, another $r(t)$ value may be computed, creating another element of a cross-correlation sequence between the incoming signal and the template.

From a computational perspective we need to do n multiplications and sum the results for each sample of the incoming signal. We either need n multipliers running in parallel or to speed up the clock with a factor of n . Then we need to figure out an efficient summing operation, in the simplest form requiring another n iterations following the multiplications. No wonder alternative pattern-matching methods are sought when power is precious.

The internal signal representation called *bitstream* is found in delta-sigma, or sigma-delta, converters, but is usually decimated to Nyquist rate binary coded numbers. However, some signal processing like filtering using bitstreams are reported [5], [6], [7], [8]. Another application of bitstreams is found in the audio coding format named Direct-Stream-Digital used in SACD, developed by Sony and Philips [9]. In this paper we explore the idea proposed in [10] of implementing cross-correlation by processing bitstreams. By doing so, quite power efficient single chip heartbeat detectors embedded in the ECG electrode are feasible.

A. Bitstream Operations

Many bitstream operations can be carried out using basic logic gates and therefore have a significant advantage compared to their multibit counterparts. Multiplication usually requiring complex digital circuitry can be done with a single XNOR gate, given a delta-sigma modulator with an input range of $[-1, 1]$. The probability that the output of the modulator, X , is 1 or 0 is then given by $P(X = 1) = (1+x)/2$ or $P(X = 0) = 1 - P(X = 1) = (1-x)/2$. The modulation of two input signals x and y is regarded as uncorrelated and

results in two bitstreams, X and Y . The XNOR of the two bitstreams, $Z = X \oplus Y$, has this property:

$$\begin{aligned} P(Z = 1) &= P(X \oplus Y = 1) \\ &= P(X = 0) \cdot P(Y = 0) + P(X = 1) \cdot P(Y = 1) \\ &= (1 + xy)/2 \end{aligned}$$

This means that the output of the XNOR has the same probability of being '1' as a delta-sigma modulation of the signal $x \cdot y$, which suggests that the simple XNOR operation is in fact a good approximation of the otherwise more complex operation of multiplying two multibit signals. The results from bitstream multiplications are just a single bit from each multiplier. A summing operation still remains before the cross-correlation result is complete.

Single gate operation allows the multiplication of the two bitstreams to be done in parallel, leaving a long sequence for summing. For power efficiency we avoid clocks exceeding the oversampled clock frequency. Knowing the multiplier results are all single bits, the summing is reduced to counting the number of '1's after multiplications. The straight forward shift-out-and-count would require increased clock speed by the length of the shift register. This register is typically more than $1k$ bits long, giving at least three orders of magnitude increase of clock speed. A novel asynchronous counter design is proposed completing the counting operation in just one, single clock cycle, maintaining the clock frequency at the oversampling rate. The counting operation is split in two pipelined sub-operations:

- 1) Sorting bit sequence from multipliers. The sorting method used is inspired by the software bubble sort algorithm, but is implemented completely asynchronously. In contrast to software bubble sort, this implementation sorts at several places at the same time. The specially designed register with sorting capabilities is named "bubble"-register.
- 2) Encoding sorted result as binary number. The sorted bit sequence is a thermometer coded representation of the result and is encoded to a binary value for further processing.

B. Bitstream Cross-Correlation

The cross-correlation is computed directly on bitstream coded signals. During setup phase, the template is shifted in directly as a bitstream coded sequence of up to 1024 bits in a template register. Then the incoming bitstream signal is shifted through the correlation register. All bits in the two registers are "multiplied" by XNOR gates at the start of every clock cycle. The bubble register is loaded with the results from the XNOR operation after an adequate delay. Then the "bubbling" is started and the rest of the clock cycle is reserved for the asynchronous sorting operation followed by latching of results. During the next clock cycle, the thermometer coded result of the bubbling is converted to a binary representation in parallel with computation of the following correlation result. It should be noted that cross-correlator results are produced at

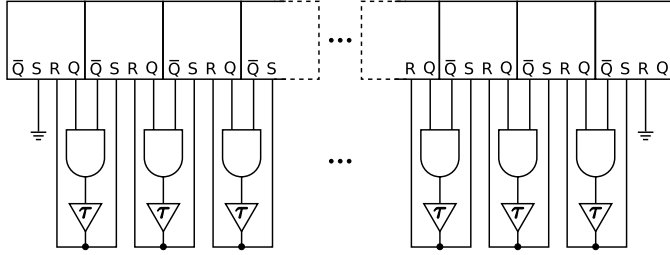


Fig. 2. Asynchronous bubble-register.

the oversampling clock rate and low-pass filtering is required for removal of the noise-shaped quantization noise inherent in the signal.

IV. IMPLEMENTATION

In these dedicated microelectronics, register lengths are hard-coded by design. Depending on application, correlation window lengths must be adapted for minimal power consumption. For easy generation of different cross-correlators we have used SKILL, a LISP-like CAD system extension language. The produced SKILL scripts facilitate fast and easy generation of both schematics and layout of cross-correlators of different sizes.

A. Bitstream Cross-Correlator

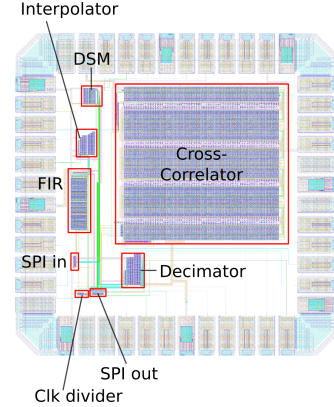
The bubble register design is shown in Fig. 2, where each bit is composed of one ordinary RS-latch, one AND gate and inverters used as delay-elements. The latches are loaded with the result from the bitstream multiplication in the beginning of the clock cycle. The bubble sorting is initiated after a predefined delay for proper settling of the latches. The final and stable condition of the bubble register has all '1's stacked to the right and all '0's to the left. This result is known as a thermometer code, and it will contain one and only one (0, 1) sequence, which may be used for unique binary encoding.

The thermometer code is converted to binary coding in the following clock cycle, while the next cross-correlation result is computed (i.e. pipelining). The thermometer coded result is fed to the binary encoder assuming a single (0, 1) transition. This transition is identified by comparing two consecutive values from the thermometer coded array using an XOR-gate. The correct precharged output lines are pulled down which encodes the desired binary result. The output lines are sampled after a predefined delay and clocked out through an SPI-inspired interface.

The calculation of one output value of the cross-correlation between two bitstream-coded signals, i.e. multiply and sum, is carried out in two clock cycles with a pipeline depth of two. The throughput is one output value per clock cycle. So, as we can see, coding signals as bitstreams facilitates simple and very power-efficient solutions.

B. Data Conditioning

Hiding of bitstream coding is an integral part of the cross-correlator chip. An interpolation filter, a delta-sigma modulator

Fig. 3. Chip layout, 1×1 mm. Includes delta-sigma converter and 1024-bits cross-correlator. Designed in STMicroelectronics 90 nm technology.

and a decimation filter gives a seamless interface to the bitstream cross-correlator. The layout of the complete system is shown in Fig. 3.

V. MEASUREMENT RESULTS

The chip is measured using a simple microcontroller, enabling measurements of the different chip modules.

A. Beat Detection

A number of automated methods for measuring intervals and segments in the cardiac cycle have been developed over the years. The lack of standardized databases containing a sufficient large number of manually annotated heartbeats, have prevented good evaluation of the performance of such algorithms. Tremendous effort from clinicians is required to manually annotate a statistically significant set of ECG records. The QT Database (qtdb) [11] includes ECG to challenge beat detection algorithms with real-world variability. Cardiologists have manually annotated sections from each record in the database.

The multi-component based cross-correlation method for beat detection was benchmarked in [4], showing an improvement compared to other beat detection approaches. The detection of waveforms, performed by the low-power bitstream cross-correlator, is presented throughout this section.

1) *Template Creation*: Each template was created by averaging the ten first identifiable components of each type. Fig. 4 shows the QRS complex and T wave templates, accompanied with the data which they were generated from. The vertical lines show the boundaries of the template, in this measurement setup confined to a width of 64 samples or 256 ms. The templates are scaled to span the input range of the delta-sigma modulator. The templates have to be the correct length to optimally identify the target pattern. The fixed template length of 256 ms is not optimal, but close to the suggested correlation interval of 300 ms in [4].

2) *Beat Detection*: The first objective is to localize the QRS complex. The decimated chip result is plotted in Fig. 5 together with the “ideal” cross-correlation, where “ideal” is the mathematical correct cross-correlation between the ECG

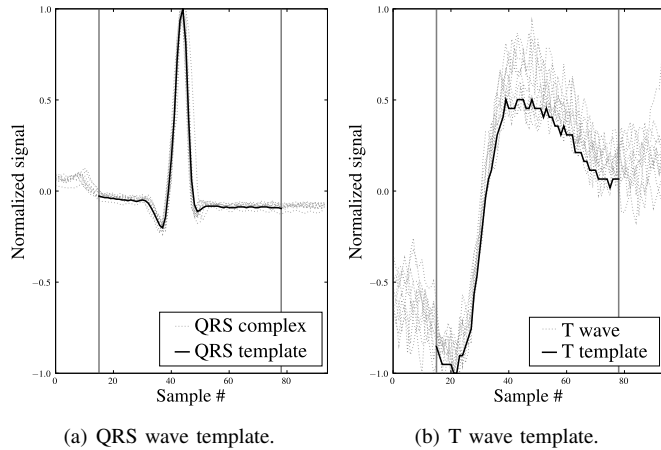


Fig. 4. Two of the used templates.

signal and the template. The “QRS” labels and the associated lines show the markers inserted by clinical experts. Solely by inspection the bitstream cross-correlation results seem promising and very close to the “ideal” result.

The chip is evaluated for detection of the QRS complex in the 30 annotated heartbeats in the record *sel100* from the qtdb. The QRS peaks were localized within a standard deviation of 3.6 ms from the expert annotations. The reduction of accuracy compared to the standard deviation of 1.8 ms reported in [4] is mostly a result of the coarse delta-sigma modulator used for producing the bitstream, and not the bitstream operation itself. The result is noisier than ideal cross-correlation, but can be improved to some extent by using a modulator giving a higher signal-to-noise ratio. Sensor data is often sampled by a delta-sigma modulator, so bitstream operations can be used to reduce processing demands without any significant loss of signal quality. This first attempt at bitstream multi-component based beat detection seems promising. A proper evaluation of the bitstream cross-correlation method for detection of components in an ECG signal demands extensive analysis, and has not yet been performed.

B. System Performance

The power consumption of the bitstream cross-correlator is proportional to the data processed. With an average load, the chip power consumption is measured to 5.1 mW performing 120 k cross-correlations each second at Nyquist rate. This is with $n_{\text{Nyquist}} = 128$, oversampling ratio = 8 and consequently $n_{\text{bitstream}} = 1024$. A rough comparison to a low power microcontroller shows a power consumption improvement of 84 times. In addition, this first chip was a proof-of-concept with far from optimal layout. The analysis was performed under the assumption that a bitstream representation of the ECG signal was already available and without decimation of the cross-correlation result.

VI. CONCLUSION

In this paper, we have presented a novel bitstream-based single-chip running cross-correlator which has been imple-

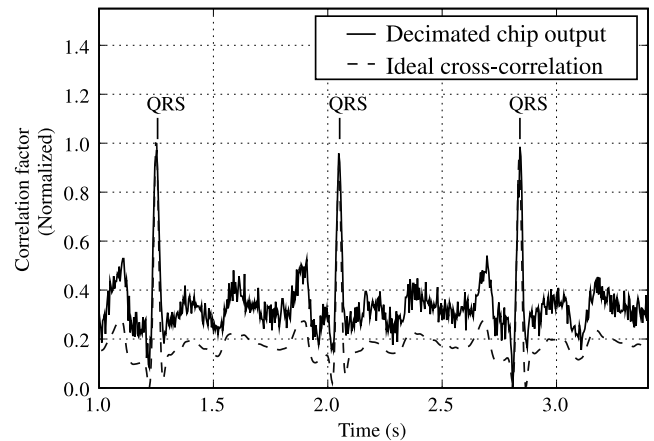


Fig. 5. Cross-correlation results for the QRS complex.

mented in 90 nm CMOS technology and which has been experimentally verified to work for ECG analysis. The low-power operation facilitates “smart” electrodes with embedded electronics for QRS-detection. The solution is generic and may well be used in other low-power applications requiring advanced pattern-matching.

ACKNOWLEDGMENT

The authors would like to thank Dept. of Informatics, University of Oslo for providing chip fabrication and lab facilities for this project.

REFERENCES

- [1] G. Yang, *Body Sensor Networks*. Springer, 2006.
- [2] B. Warneke, M. Last, B. Liebowitz, and K. Pister, “Smart dust: communicating with a cubic-millimeter computer,” *Computer*, vol. 34, no. 1, pp. 44–51, Jan 2001.
- [3] B.-U. Kohler, C. Hennig, and R. Orglmeister, “The principles of software qrs detection,” *Engineering in Medicine and Biology Magazine, IEEE*, vol. 21, no. 1, pp. 42–57, Feb 2002.
- [4] T. Last, C. Nugent, and F. Owens, “Multi-component based cross correlation beat detection in electrocardiogram analysis,” *BioMedical Engineering OnLine*, vol. 3, no. 1, p. 26, 2004. [Online]. Available: <http://www.biomedical-engineering-online.com/content/3/1/26>
- [5] P. O’Leary and F. Maloberti, “Bit stream adder for oversampling coded data,” *Electronics Letters*, vol. 26, pp. 1708–1709, Sept 1990.
- [6] F. Maloberti and P. O’Leary, “Processing of signals in their oversampled delta-sigma domain,” in *Circuits and Systems, 1991. Conference Proceedings, China., 1991 International Conference on*, Jun 1991, pp. 438–441 vol.1.
- [7] F. Maloberti, “Non conventional signal processing by the use of sigma delta technique: a tutorial introduction,” in *Circuits and System, ISCAS ’92. Proceedings, 1992 IEEE International Symposium on*, vol. 6, May 1992, pp. 2645–2648.
- [8] S. Summerfield, S. Kershaw, and M. Sandler, “Sigma-delta bitstream filtering in vlsi,” in *Circuits and Systems, 1994., Proceedings of the 37th Midwest Symposium on*, vol. 2, Aug 1994, pp. 1200–1203 vol.2.
- [9] E. Janssen and D. Reefman, “Super-audio cd: an introduction,” *Signal Processing Magazine, IEEE*, vol. 20, no. 4, pp. 83–90, July 2003.
- [10] T. Lande, T. Constandinou, A. Burdett, and C. Toumazou, “Running cross-correlation using bitstream processing,” *Electronics Letters*, vol. 43, no. 22, 2007.
- [11] P. Laguna, R. Mark, A. Goldberger, and G. Moody, “A database for the evaluation of algorithms for measurement of qt and other waveform intervals in the ecg,” *Computers in Cardiology*, vol. 24, pp. 673–676, 1997.

Paper 19 Power efficient cross-correlation using bitstreams

O. E. Liseth, D. Mo, H. A. Hjortland, T. S. Lande, and D. T. Wisland. “Power efficient cross-correlation using bitstreams”. In: *NORCHIP, 2009*, pp. 1–4, Trondheim, Nov. 2009. doi:10.1109/NORCHIP.2009.5397844. © 2009 IEEE. Reprinted with permission.

My Contributions to This Paper

As co-supervisor I worked with the students, with paper preparation and analysis of the concept. This paper presents more system aspects of a template matching chip using bitstream coding (very related to CTBV coding).

Power Efficient Cross-Correlation Using Bitstreams

Olav E. Liseth, Daniel Mo, Håkon A. Hjortland, Tor Sverre “Bassen” Lande and Dag T. Wisland

Dept. of Informatics, University of Oslo, Norway

Email: olaveli@ifi.uio.no, daniemo@student.matnat.uio.no, haakoh@ifi.uio.no, bassen@ifi.uio.no, dagwis@ifi.uio.no

Abstract—The fundamental operation of cross-correlating signals is viable in a number of signal processing applications. In typical pattern-matching applications, cross-correlation is desirable. In this paper we present a power efficient implementation of a time-domain cross-correlator suitable for integration in CMOS. Bitstream coding of both data and template simplify multiplication operations. Measured performance of a CMOS implementation in 90 nm technology is reported.

I. INTRODUCTION

A major trend in microelectronics is integration of specialized solutions in an increasing number of applications. The idea of ubiquitous computing is certainly exciting, but at the same time demanding. Both sensing and controlling real world processes demand mixed-mode solutions combined with challenging signal conditioning and processing. The notion of small, portable, battery-operated systems often organized in a wireless sensor network (WSN) has initiated significant research activity. In these applications size is limited and low-power operation is mandatory for battery operation. The benefit of adopting specialized silicon systems is evident in applications like WSN motes [1].

An overall characteristic of these ubiquitous computational devices is mixed-mode operation. Sensing of external states is accomplished with analog-to-digital converters (ADCs) and controlling of external processes requires DACs. A popular and power-efficient converter architecture is sigma-delta, or delta-sigma, converters well suited for integration in digital technology.

A recurring signal processing task in real-world signal analysis is pattern-matching recovering special features of some sensed signal. As an example in this work we will use signal processing for Electrocardiogram (ECG) classification. Although filter-based solutions have been developed with great sophistication over the last decades [2], recent publications [3] indicate that cross-correlation based methods are preferable. Furthermore, the miniaturization of ECG-monitoring devices are pursued both in research and in industry [4]. The notion of heartbeat detectors embedded in the ECG electrode and configured in a wireless sensor network is tempting and would enable long-term ECG analysis. Substitution of the quite clumsy Holter monitors used today would make life easier.

In this work we will show how power-efficient cross-correlators are implementable in standard CMOS technology, exploring bitstream-coded signals. The internal signal representation called *bitstream* is found in sigma-delta converters, but is usually decimated to Nyquist rate binary coded numbers. However, some signal processing like filtering using bitstreams

are reported [5], [6], [7], [8]. Another application of bitstreams is found in the audio coding format named Direct-Stream-Digital used in SACD, developed by Sony and Philips [9].

The idea of cross-correlation using bitstreams was proposed in [10] and evaluated for heart rate variability study in [11]. In this paper we present an implementation of a complete cross-correlation chip with a binary coded interface.

II. BITSTREAM CROSS-CORRELATOR

A discrete estimate of the cross-correlation between two sequences is found by the equation:

$$r(t) = \sum_{k=0}^{n-1} y(k)x(t+k)$$

The sequence $x(t)$ is cross-correlated with a template sequence, $y(t)$, by multiplying each element of the two sequences over a window of length n and summing the result. The result, $r(t)$, is a good estimate of the cross-correlation between the two sequences. For every new sample of the incoming signal, another $r(t)$ value may be computed, creating another element of a cross-correlation sequence between the incoming signal and the template.

From a computational perspective we need to do n multiplications and sum the results for each sample of the incoming signal. We either need n multipliers running in parallel or to speed up the clock with a factor of n . Then we need to figure out an efficient summing operation, in the simplest form requiring another n iterations following the multiplications. No wonder alternative pattern-matching methods are sought when power is precious.

In this paper we explore the idea proposed in [10] of implementing cross-correlation by processing bitstreams. By doing so, quite power efficient single chip heartbeat detectors embedded in the ECG electrode are feasible.

III. SINGLE-CHIP CROSS-CORRELATOR

To allow for a simple interface to the chip, both input and output signals are assumed to be Nyquist rate binary encoded. Conversion to and from oversampled bitstream representation is done on-chip. Fig. 1 shows how the bitstream is only necessary internal to the system. However, a Serial Peripheral Interface (SPI) bus is used for interfacing, and on-chip multiplexers facilitate testing of individual blocks.

A. Bitstream Conversion

The binary-to-bitstream signal conversion is done using a two step interpolation filter and a sigma-delta modulator. The interpolation filter upsamples an input signal at the Nyquist rate by the oversampling ratio required by the modulator. Interpolation in several steps is commonly used. This allows for a combination of an anti-alias filter, with a narrow transition band, and a more hardware efficient Cascaded Integrator Comb (CIC) filter [12]. This is a compromise between filter quality and silicon area, which in turn affects power consumption.

Similarly, a decimation function is required for removing the high-frequency quantization noise present in bitstream coded signals, shown as the CIC decimator in Fig. 1.

Using a low oversampling ratio (OSR) when modulating the bitstream obtains the largest power savings [10]. As a consequence, the possible signal quality of the bitstream is limited. Compromising between these two considerations, the system OSR is set to 8. Still the bit sequence may be of significant length depending on the time-span of the correlation window. In our test-chip we use a correlation length of 1024 bits.

B. Bitstream Operations

Many bitstream operations can be carried out using basic logic gates and therefore have a significant advantage compared to their multibit counterparts. Multiplication usually requiring complex digital circuitry can be done with a single XNOR gate, given a delta-sigma modulator with an input range of $[-1, 1]$. The probability that the output of the modulator, X , is 1 or 0 is then given by $P(X = 1) = (1+x)/2$ or $P(X = 0) = 1 - P(X = 1) = (1-x)/2$. The modulation of two input signals x and y is regarded as uncorrelated and results in two bitstreams, X and Y . The XNOR of the two bitstreams results in:

$$\begin{aligned} P(X \oplus Y = 1) &= P(X = 0) \cdot P(Y = 0) \\ &\quad + P(X = 1) \cdot P(Y = 1) \\ &= (1 + xy)/2 \end{aligned}$$

which indicates that operations usually requiring complex digital circuitry can be done with a simple logic gate when processing bitstreams. The results from bitstream multiplications are just a single bit from each multiplier. A summing operation still remains in order to calculate the cross-correlation value.

For power efficiency we try to avoid clocks exceeding the oversampled clock frequency. Knowing the multiplier results are all single bits, the summing is reduced to counting the number of '1's after multiplications. A novel asynchronous counter design is proposed, completing the counting operation in just one single clock cycle, maintaining the clock frequency at the oversampling rate. The counting operation is split in two pipelined sub-operations:

- 1) Sorting bit sequence from multipliers. The sorting method used is inspired by the software bubble sort algorithm, but is implemented completely asynchronously. In contrast to software bubble sort, this implementation sorts at several places at the same time. The specially

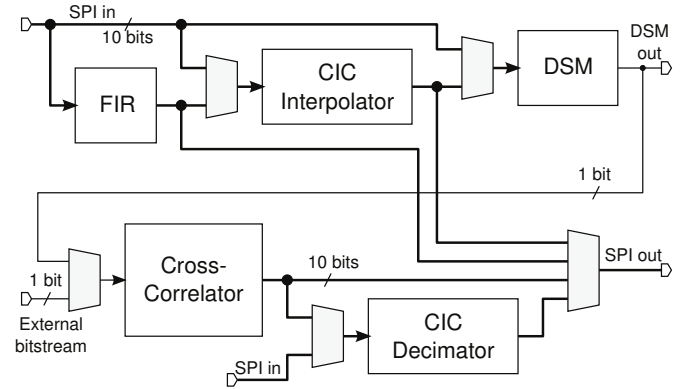


Fig. 1. Block schematic showing main signal paths. The multiplexers allow each block to be tested individually. During normal operation, the only IO pins required are three pins for the SPI-inspired interface.

designed register with sorting capabilities is named “bubble” register.

- 2) Encoding sorted result as binary number. The sorted bit sequence is a thermometer coded representation of the result and is encoded to a binary value for further processing.

The asynchronous operation is achieved using inherent gate delays explained below.

C. Bitstream Cross-Correlation

During setup phase, the binary-to-bitstream converter may be used or the template may be shifted in directly as a bitstream coded sequence of up to 1024 bits. Then the incoming signal is shifted into the correlation register coded as a bitstream or converted by the binary-to-bitstream converter. All bits in the two registers are “multiplied” by XNOR gates at the start of every clock cycle. The bubble register is loaded with the results from the XNOR operation after an adequate delay. Then the “bubbling” is started and the rest of the clock cycle is reserved for the asynchronous sorting operation followed by latching of results. During the next clock cycle the thermometer coded result of the bubbling is converted to a binary representation in parallel with computation of the following correlation result.

IV. IMPLEMENTATION

In these dedicated systems, register lengths are hard-coded by design. Depending on application, correlation window lengths must be adapted for minimal power consumption. For easy generation of different cross-correlators we have used SKILL, a LISP-like CAD system extension language. The produced SKILL scripts facilitate fast and easy generation of both schematics and layout of cross-correlators of different sizes.

A. Bitstream Cross-Correlator

The bubble register used for sorting is shown in Fig. 2, where each element consists of one ordinary RS-latch, one

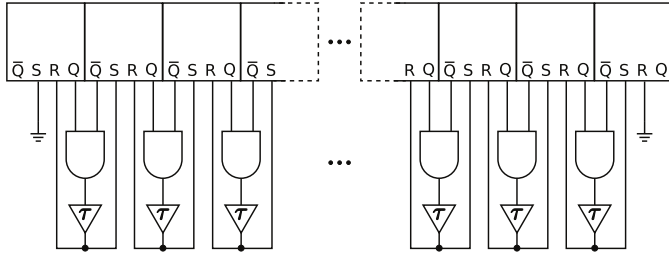


Fig. 2. Asynchronous bubble register.

AND gate and inverters used as delay-elements. The latches are loaded with the result from the bitstream multiplication in the beginning of the clock cycle. The bubble sorting is initiated after a predefined delay for proper settling of the latches.

It is important to notice that the exchange of bits in the bubble register is local, enabling parallel operation. Basically, the bubble operation $(1, 0) \rightarrow (0, 1)$ will be carried out a short delay after the input reaches this state. To ensure safe operation the actual exchange operation is delayed using some inverters. In this way all the '1's are "bubbled" to the right while the '0's are "bubbled" to the left.

The final and stable condition of the bubble register has all '1's stacked to the right and all '0's to the left. This result is known as a thermometer code, and it will contain one and only one $(0, 1)$ sequence, which may be used for unique binary encoding.

The thermometer code is converted to binary coding in the following clock cycle, while the next cross-correlation result is computed. The thermometer coded result is fed to the binary encoder assuming a single $(0, 1)$ transition. This transition is identified by comparing two consecutive values from the thermometer coded array using an XOR-gate. The correct precharged output lines are pulled down which encodes the desired binary result. The output lines are sampled after a predefined delay and clocked out through an SPI-inspired interface.

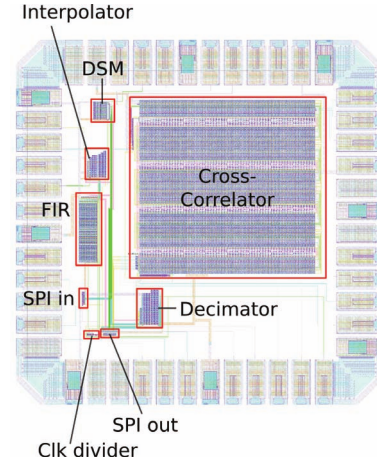
B. Data Conditioning

Hiding of bitstream coding is an integral part of the cross-correlator chip. The following modules are included:

1) *Interpolation Filter*: The anti-aliasing step of the interpolation filter is a 20-tap halfband FIR filter following an upsampling by two. This filter type has every other coefficient set to zero and is an efficient way to achieve a narrow transition band around 0.5π . Each tap in the filter is in turn simplified. Gain steps are quantized such that each are realizable as a sum of maximum two hardwired bit-shift operations. This removes the requirements for full multiplications in the filter.

The CIC filter step is of third order and performs an upsampling by four, resulting in the desired oversampling rate.

2) *Sigma-Delta Modulator*: The modulator has a second order error feedback structure. This structure uses multibit feedback and is well suited for digital input. The most important design constraint for the modulator is the requirement that the output is a bitstream. This rules out good modulator types

Fig. 3. Chip layout, 1×1 mm. Includes delta-sigma converter and 1024-bits cross-correlator. Designed in STMicroelectronics 90 nm technology.

such as cascaded or MASH architectures which yield word streams. Quality of the modulator is also limited by the low OSR in the system. For an OSR of 8, modulator orders of two and more give a maximum expected Signal to Quantization Noise Ratio (SQNR) of 35–40 dB.

3) *Decimation Filter*: The filter used for decimation is an ordinary CIC decimation filter of third order with a downsampling ratio of 8. This converts the oversampled output from the cross-correlation block back to a Nyquist rate signal, while reducing high frequency noise inherited from the bitstream representation.

V. MEASUREMENT RESULTS

The chip is measured using a simple microcontroller, enabling measurements of the different chip modules.

1) *Interpolation Filter*: The frequency response of the cascaded interpolation filter is shown in Fig. 4. The response of the FIR filter step was confirmed by chip measurements, using both white noise input signals and sinusoidal inputs. Total stop band attenuation is 35–40 dB, close to expected performance.

2) *Sigma-Delta Modulator*: The chip modulator was tested using a full-scale sinusoidal input. Output spectrum is shown in Fig. 5 and shows the expected SQNR level.

A. System Performance

Here we briefly demonstrate the capabilities of the system by performing component-based ECG-analysis [3]. We also estimate the performance figures of the implementation.

In Fig. 6 the cross-correlation from the test-chip is shown together with cross-correlation using the *xcorr()*-function in MATLAB. The data used is taken from the QT database [13]. The first QRS-wave from the annotated database is selected as template and loaded into the template-register as a bitstream. Then a sequence of three heartbeats is fed to the chip and the cross-correlation result is decimated and plotted. Just by inspection the cross-correlation results are promising and very close to results obtained using "ideal" cross-correlation with MATLAB. A proper evaluation of the cross-correlation chip

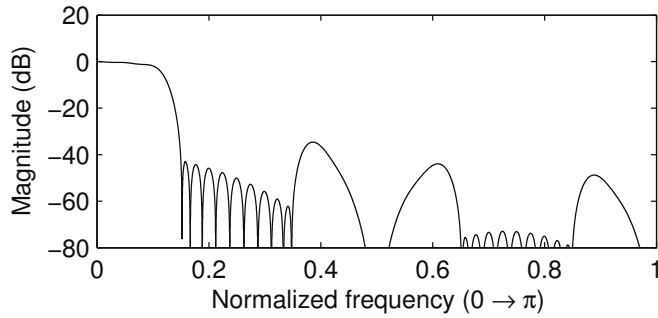


Fig. 4. Frequency response of cascaded interpolation filter. Normalized to half the oversampling frequency.

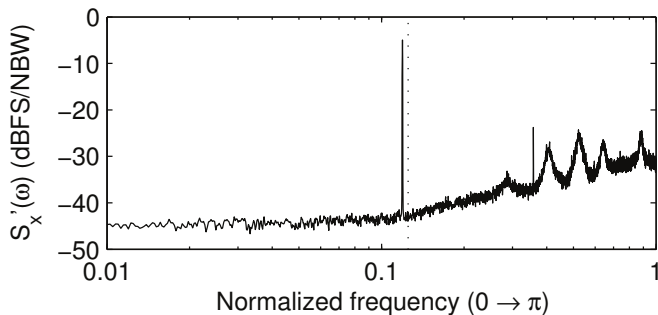


Fig. 5. Power spectral density of the sigma-delta modulator. Nyquist frequency indicated by dotted line.

for QRS-detection demands extensive analysis, far beyond the scope of this paper.

The main computational block doing cross-correlation is measured to 2.1 mW power consumption when active. With a clock frequency of 480 kHz and an OSR of 8 this is equivalent to approx. 7.7 M multiplications each second at Nyquist rate. In addition, this first chip was a proof-of-concept with far from optimal layout. We consider these results to be very promising for low-power, single-chip pattern-recognition.

As indicated in the introduction, cross-correlation is a generic pattern-matching computational element suited for several signal processing tasks. The bitstream processing solution presented in this paper aims at low-power operation at low or moderate signal frequencies. There is a growing demand for low frequency sensor interfacing where filtering is required. The proposed cross-correlator chip may also be used as a programmable filter by time-warping the template for convolution. The low-frequency operation and programmable filter function can be useful in many emerging biomedical applications.

VI. CONCLUSION

In this paper we have presented a novel single-chip cross-correlator suitable for power-efficient pattern matching. Bitstream encoded signals found in sigma-delta converters are used for efficient multiply-and-sum operations basically substituting multipliers with simple gates. A novel asynchronous bubble register is utilized avoiding increased clock frequency.

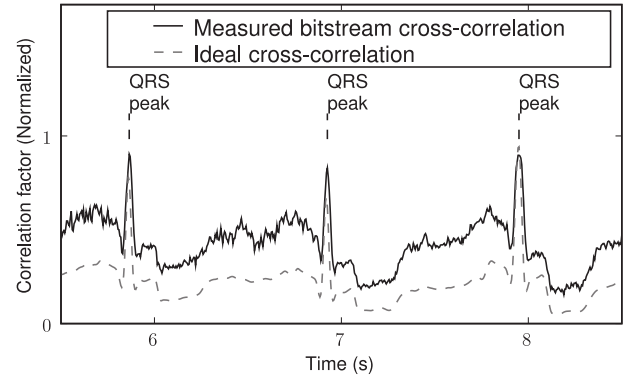


Fig. 6. Bitstream cross-correlation results.

The running cross-correlator is implemented in 90 nm CMOS and measured results are provided. We expect these kind of cross-correlators to be viable in power-limited, low signal frequency applications like QRS-detection in ECG-signals.

ACKNOWLEDGMENT

The authors would like to thank Dept. of Informatics, University of Oslo for providing chip fabrication and lab facilities for this project.

REFERENCES

- [1] B. Warneke, M. Last, B. Liebowitz, and K. Pister, "Smart dust: communicating with a cubic-millimeter computer," *Computer*, vol. 34, no. 1, pp. 44–51, Jan 2001.
- [2] [Online]. Available: <http://www.openecg.org/>
- [3] T. Last, C. Nugent, and F. Owens, "Multi-component based cross correlation beat detection in electrocardiogram analysis," *BioMedical Engineering OnLine*, vol. 3, no. 1, p. 26, 2004. [Online]. Available: <http://www.biomedical-engineering-online.com/content/3/1/26>
- [4] [Online]. Available: <http://www.toumaz.com/>
- [5] P. O'Leary and F. Maloberti, "Bit stream adder for oversampling coded data," *Electronics Letters*, vol. 26, pp. 1708–1709, Sept 1990.
- [6] F. Maloberti and P. O'Leary, "Processing of signals in their oversampled delta-sigma domain," in *Circuits and Systems, 1991. Conference Proceedings, China, 1991 International Conference on*, Jun 1991, pp. 438–441 vol.1.
- [7] F. Maloberti, "Non conventional signal processing by the use of sigma delta technique: a tutorial introduction," in *Circuits and System, ISCAS '92. Proceedings, 1992 IEEE International Symposium on*, vol. 6, May 1992, pp. 2645–2648.
- [8] S. Summerfield, S. Kershaw, and M. Sandler, "Sigma-delta bitstream filtering in vlsi," in *Circuits and Systems, 1994., Proceedings of the 37th Midwest Symposium on*, vol. 2, Aug 1994, pp. 1200–1203 vol.2.
- [9] E. Janssen and D. Reefman, "Super-audio cd: an introduction," *Signal Processing Magazine, IEEE*, vol. 20, no. 4, pp. 83–90, July 2003.
- [10] T. Lande, T. Constandinou, A. Burdett, and C. Toumazou, "Running cross-correlation using bitstream processing," *Electronics Letters*, vol. 43, no. 22, pp. 25, 2007.
- [11] M. Ho, T. Lande, and C. Toumazou, "Efficient computation of the lf/hf ratio in heart rate variability analysis based on bitstream filtering," in *Biomedical Circuits and Systems Conference*, Nov 2007, pp. 17–20.
- [12] E. B. Hogenauer, "An economical class of digital filters for decimation and interpolation," *Acoustics, Speech, and Signal Processing, IEEE Transactions on*, vol. 29, no. 2, pp. 155–162, Apr 1981.
- [13] P. Laguna, R. Mark, A. Goldberger, and G. Moody, "A database for the evaluation of algorithms for measurement of qt and other waveform intervals in the ecg," *Computers in Cardiology*, vol. 24, pp. 673–676, 1997.

Paper 20 Power-Efficient Cross-Correlation Beat Detection in Electrocardiogram Analysis Using Bitstreams

O. E. Liseth, D. Mo, H. A. Hjortland, T. S. Lande, and D. T. Wisland. “Power-Efficient Cross-Correlation Beat Detection in Electrocardiogram Analysis Using Bitstreams”. *Biomedical Circuits and Systems, IEEE Transactions on*, Vol. 4, No. 6, pp. 419–425, Dec. 2010. doi : 10.1109/TBCAS.2010.2079933. © 2010 IEEE. Reprinted with permission.

My Contributions to This Paper

As co-supervisor of these MS students I worked with them during development and implementation of a running cross correlator. Somewhat involved in the actual writing as well.

Power-Efficient Cross-Correlation Beat Detection in Electrocardiogram Analysis Using Bitstreams

Olav E. Liseth, Daniel Mo, Håkon A. Hjortland, *Member, IEEE*, Tor Sverre “Bassen” Lande, *Fellow, IEEE*, and Dag T. Wisland, *Member, IEEE*

Abstract—In this paper, we present a novel cross-correlator chip suitable for “smart” ECG electrodes. Sophisticated QRS-detection is feasible exploring multicomponent-based cross-correlation and by exploiting the simplicity offered by bitstream processing. The chip is evaluated using real ECG signals as an example application. Power-efficient running cross correlation is obtained with novel asynchronous circuit solutions consuming approximately 730 μ W while processing ECG data sampled at a rate of 250 Hz. The implemented generic cross-correlator may be explored for other pattern-matching applications in wireless sensor networks or other low-power applications.

Index Terms—Biomedical computing, bitstream operations, complementary metal-oxide semiconductor (CMOS), cross-correlation, sigma-delta modulation.

I. INTRODUCTION

THE NOTION of small, portable, battery-operated systems often organized in a wireless sensor network (WSN) has initiated significant research activity. Usage of WSNs in body sensor networks (BSN) [1] is interesting and may facilitate elegant and compact solutions for vital sign monitoring. In BSN applications, size is limited and low-power operation is mandatory for battery operation. The benefit of adopting specialized silicon systems for minimal size and power consumption in BSN applications is evident [2].

An interesting application of BSNs is personalized health monitoring. Small sensor nodes may be attached to the body and even located inside the body as implants. Organized as WSNs, health information is processed, conditioned and wirelessly interacting with surrounding networks.

An important vital sign indicator is the heartbeat or the more accurate analysis of the electrical heart activity known as ECG analysis. Measuring of heartbeats of the simplest form may be done by sensing an artery’s pulsation at the body surface and is merely usable for heart rate measurements. The elaborate ECG-analysis requires significant bedside equipment and typically 12 electrodes attached to the body. These procedures are often carried out in hospitals.

However, there is an intermediate ECG analysis often used for long-term observations known as Holter monitors. These

systems tend to have fewer electrodes, normally 2–3 attached to the torso. The Holter monitor is worn during normal activity and should not interfere with normal lifestyle. These monitors do catch serious cardiac arrhythmia like Bradycardia, Tachycardia, Sinus arrest, Ventricular tachycardia, Supraventricular tachycardia and similar. Older systems left hours of ECG signals to be analyzed by cardiologists, but over the years a number of methods have been proposed for computerized ECG analysis [3].

In this paper we are proposing a novel single-chip cross-correlator which can be used to perform ECG analyses. A programmable “smart” ECG electrode with embedded heart-beat detection or even more elaborate ECG analyses is feasible. Combined with body area network technology, a wireless ECG system for long-term usage may be constructed.

II. HEARTBEAT DETECTION

Beat detection involves identifying all cardiac cycles in ECG recordings and locating each identifiable waveform component within a cycle, the P wave peak, the QRS peak and the T wave peak, including the associated on- and offset timing. The electrical activity from one cycle of cardiac contraction and relaxation is schematically drawn in Fig. 1. The accuracy of the beat detection has significant impact on the overall classification performance and various computerized methods have been developed over the years trading off computational efficiency with detection quality.

A. Multicomponent-Based Heartbeat Detection

Although cross-correlation has been explored for template-matching ECG-signals with anomaly templates, this method was abandoned due to both poor performance and high computational demands. Recently an improvement of the traditional cross-correlation method was proposed in [4], where each of the identifiable waveform components within a cycle are searched for in isolation. Three templates were used, one for the P wave, one for the QRS complex and one for the T wave. The first step is to locate the QRS complex by cross-correlating the QRS template with the ECG signal. This method is repeated with the P and the T wave templates. In each case, the highest value of the cross-correlation result within the given correlation interval is compared to a predefined threshold. The threshold value is established during a pre-learning phase and can be adjusted during operation. Although the component-based heartbeat detection is functionally promising, the computational complexity of cross-correlation is still demanding. In WSN applications, power-efficient implementations are required.

Manuscript received May 14, 2010; revised August 18, 2010; accepted September 14, 2010. Date of publication October 28, 2010; date of current version November 24, 2010. This paper was recommended by Associate Editor G. Yuan.

The authors are with the Department of Informatics, University of Oslo, Oslo N-0316, Norway (e-mail: olaveli@ifi.uio.no).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TBCAS.2010.2079933

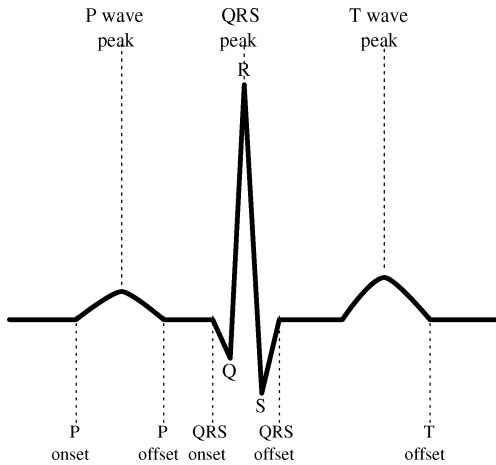


Fig. 1. Schematic representation of a normal ECG waveform including P wave, QRS complex, and T wave beat markers.

III. SINGLE-CHIP CROSS-CORRELATOR

The discrete cross-correlation between two sequences is found by the equation:

$$(x \star y)[t] = \sum_{k=0}^{n-1} x(k)y(t+k)$$

The template sequence, $x(t)$, is cross-correlated with the input signal, $y(t)$, by multiplying corresponding elements from the two sequences over a window of length n , given by the template length, and summing the result. For every new sample of the incoming signal, another value may be computed, creating another element of a cross-correlation sequence between the incoming template and the signal.

From a computational perspective, we need to do n multiplications and sum the results for each sample of the incoming signal. We either need n multipliers running in parallel or to speed up the clock with a factor of n . Then we need to figure out an efficient summing operation, in the simplest form requiring another n iterations following the multiplications. This makes cross-correlation a heavy calculation, which makes it understandable that alternative pattern-matching methods are sought when power is precious.

A. Bitstream Representation

The internal signal representation called *bitstream* is found in delta-sigma, or sigma-delta, converters, but is usually decimated to Nyquist rate binary coded numbers. However, some signal processing using bitstreams are reported [5]–[8]. Another application of bitstreams is found in the audio coding format named Direct-Stream-Digital used in Super Audio CD (SACD), developed by Sony and Philips [9]. In this paper, we explore the idea proposed in [10] of performing cross-correlation by processing bitstreams. By doing so, quite power-efficient single chip heart-beat detectors embedded in the ECG electrode are feasible.

B. Bitstream Conversion

The binary-to-bitstream signal conversion is done using a two step interpolation filter and a sigma-delta modulator. The interpolation filter upsamples an input signal at the Nyquist rate by

the oversampling ratio required by the modulator. Interpolation in several steps is commonly used. This allows for a combination of an anti-alias filter, with a narrow transition band, and a more hardware-efficient cascaded integrator comb (CIC) filter [11]. This is a compromise between filter quality and silicon area which, in turn, affects power consumption.

Similarly, a decimation function is required for removing the high-frequency quantization noise present in bitstream coded signals and to downsample the signal to Nyquist rate.

Using a low oversampling ratio (OSR) when modulating the bitstream obtains the largest power savings [10]. As a consequence, the possible signal quality of the bitstream is limited. Compromising between these two considerations, the system OSR is set to 8. Still the bit sequence may be of significant length depending on the time span of the correlation window. In our test chip we use a template length of 1024 bits, limited to 512 bits due to a hardware bug. In this prototype implementation we have not taken any measures to prevent pattern noise in the bitstream encoding. In a more advanced implementation techniques like dithering might be added to the bitstream encoder.

C. Bitstream Operations

Many bitstream operations can be carried out using basic logic gates and, therefore, have a significant advantage compared to their multibit counterparts. Multiplication usually requiring complex digital circuitry can be done with a single XNOR gate, given a delta-sigma modulator with an input range of $[-1, 1]$. The probability that the output of the modulator, X , is 1 or 0 is then given by $P(X = 1) = (1 + x)/2$ or $P(X = 0) = 1 - P(X = 1) = (1 - x)/2$. The modulation of two input signals x and y is regarded as uncorrelated and results in two bitstreams, X and Y . The XNOR of the two bitstreams $Z = X \oplus Y$ has this property

$$\begin{aligned} P(Z = 1) &= P(X \oplus Y = 1) \\ &= P(X = 0) \cdot P(Y = 0) \\ &\quad + P(X = 1) \cdot P(Y = 1) \\ &= (1 + xy)/2. \end{aligned}$$

This means that the output of the XNOR has the same probability of being “1” as a delta-sigma modulation of the signal $x \cdot y$, which suggests that the simple XNOR operation is, in fact, a good approximation of the otherwise more complex operation of multiplying two multibit signals.

Single gate operation allows the multiplication of the two bitstreams to be done in parallel, leaving a long sequence for summing. For power efficiency we avoid clocks exceeding the oversampled clock frequency. Knowing the multiplier results are all single bits, the summing is reduced to counting the number of “1”s after multiplications. The straight forward shift-out-and-count would require the clock speed to be increased by a factor corresponding to the length of the shift register. This register is typically more than 1 kbits long, giving at least three orders of magnitude increase of clock speed. A novel asynchronous counter design is proposed, completing the counting operation in just one single clock cycle, maintaining the clock frequency at the oversampling rate. The counting operation is split in two pipelined suboperations as follows.

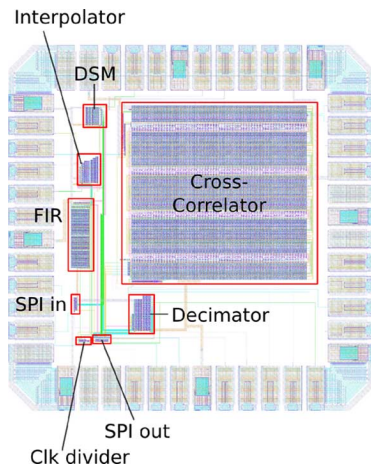


Fig. 2. Chip layout, 1×1 mm. Includes delta-sigma converter and 1024-bits cross-correlator. Designed in STMicroelectronics 90-nm technology.

- 1) Sorting bit sequence from multipliers. The sorting method used is inspired by the software bubble sort algorithm, but is implemented completely asynchronously. In contrast to software bubble sort, this implementation sorts at several places at the same time. The specially designed register with sorting capabilities is named “bubble” register.
- 2) Encoding sorted result as a binary number. The sorted bit sequence is a thermometer coded representation of the result and is encoded to a binary value for further processing.

D. Bitstream Cross-Correlation

The cross-correlation is computed directly on bitstream coded signals. During setup phase, the template is shifted in directly as a bitstream coded sequence of up to 1024 bits in a template register. Then the incoming bitstream signal is shifted through the correlation register. All bits in the two registers are “multiplied” by XNOR gates at the start of every clock cycle. The bubble register is loaded with the results from the XNOR operation after an adequate delay. Then the “bubbling” is started and the rest of the clock cycle is reserved for the asynchronous sorting operation followed by latching of results. During the next clock cycle, the thermometer coded result of the bubbling is converted to a binary representation in parallel with computation of the following correlation result. It should be noted that cross-correlator results are produced at the oversampling clock rate and low-pass filtering is required for removal of the noise-shaped quantization noise inherent in the signal.

IV. IMPLEMENTATION

In these dedicated microelectronics, register lengths are hard-coded by design. Depending on application, correlation window lengths must be adapted for minimal power consumption. For easy generation of different cross-correlators we have used SKILL, a LISP-like CAD system extension language. The produced SKILL scripts facilitate fast and easy generation of both schematics and layout of cross-correlators of different sizes. The layout of the complete system is shown in Fig. 2.

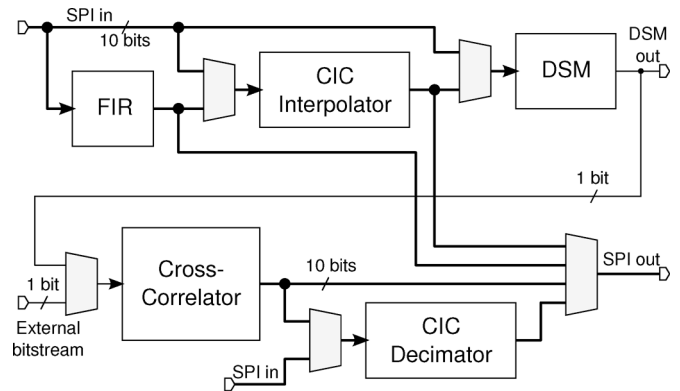


Fig. 3. Block diagram of the implemented chip showing main signal paths.

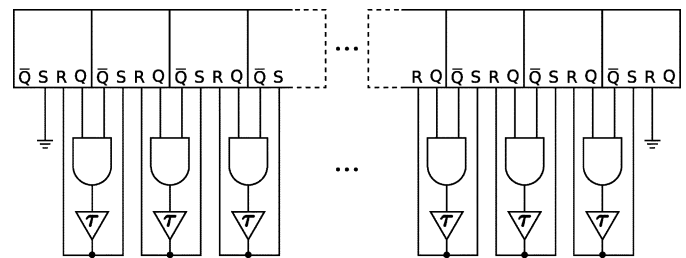


Fig. 4. Asynchronous bubble register.

To allow for a simple interface to the chip, both input and output signals are assumed to be Nyquist rate binary encoded. Conversion to and from oversampled bitstream representation is done on-chip. Fig. 3 shows how the bitstream is only necessary internally in the system. A serial peripheral interface (SPI)-inspired bus is used for interfacing, and on-chip multiplexers facilitate testing of individual blocks. This also allows for a choice between operation of the full signal chain or only the cross-correlator block. During normal operation, the only IO pins required are four pins for the SPI-inspired interface. When the cross-correlator is operating alone, the input is a bitstream and output is words over SPI. This mode also requires more off-chip processing.

A. Bitstream Cross-Correlator

The bubble register design is shown in Fig. 4, where each bit is composed of one ordinary RS-latch, one AND gate and inverters used as delay-elements. The latches are loaded with the result from the bitstream multiplication in the beginning of the clock cycle. The bubble sorting is initiated after a predefined delay for proper settling of the latches. The final and stable condition of the bubble register has all “1”s stacked to the right and all “0”s to the left. This result is known as a thermometer code, and it will contain a single (0, 1) sequence, which may be used for unique binary encoding.

The thermometer code is converted to binary coding in the following clock cycle, while the next cross-correlation result is computed (i.e., pipelining). The thermometer coded result is fed to the binary encoder assuming a single (0, 1) transition. This transition is identified by comparing two consecutive values from the thermometer coded array using an XOR-gate. The correct precharged output lines are pulled down which encodes the

desired binary result and the output lines are sampled after a pre-defined delay. The result can optionally be decimated on-chip before being read out through the SPI-inspired interface.

The calculation of one output value of the cross-correlation between two bitstream-coded signals (i.e., multiply and sum) is carried out in two clock cycles with a pipeline depth of two. The throughput is one output value per clock cycle. So as we can see, coding signals as bitstreams facilitates simple and very power-efficient solutions.

1) *Transistor Count*: This first chip was a proof-of-concept with far from optimal design. Buffer sizes and time delays were exaggerated, debug logic and extra latch stages were added to ensure expected behavior and all logic was built up solely from NAND and inverter cells. Based on vendor provided cell library specifications, both power and size could be reduced significantly by using those cells instead.

The bitstream cross-correlator consists of two signal shift registers, XNOR gates for the bitstream multiplication and the bubble register. This is implemented using 190 transistors per bit. A redesign with vendor library cells and without debugging logic would reduce the transistor count to 78 transistors per bit. A comparison of the transistor count in a standard digital approach to cross-correlation and this bitstream solution is done in [10].

The number of bubble operations required to perform one sorting in the bubble register is $O(n^2)$, where n denotes the number of bits in the register. This activity involves the elements shown in Fig. 4. A feasible implementation consisting of a NOR-based RS-latch, a NAND gate and an inverter adds up to 14 transistors per bit. The thermometer to binary encoder uses on average $\log_2(n)/2 + 8$ transistors per bit.

The bubble register is a novel solution which may be beneficial in some contexts. The fact that the bubble sort is done asynchronously is spreading the instantaneous clock load. Although there may be alternative power-efficient implementations of the counting operation, the bubble-sort register is easy to scale and represents a novel “clockless” design approach which may be explored further.

B. Data Conditioning

Hiding of bitstream coding is an integral part of the cross-correlator chip. The following modules give a seamless interface to the bitstream cross-correlator, shown in Fig. 3:

1) *Interpolation Filter*: The anti-aliasing step of the interpolation filter is a 20-tap halfband FIR filter following an up-sampling by two. This filter type has every other coefficient set to zero and is an efficient way to achieve a narrow transition band around 0.5π . Each tap in the filter is, in turn, simplified. Gain steps are quantized so that each are realizable as a sum of maximum two hardwired bit-shift operations. This removes the requirements for full multiplications in the filter.

The CIC filter step is of third order and performs an up-sampling by four, resulting in the desired oversampling rate.

2) *Sigma-Delta Modulator*: The modulator has a second order error feedback structure. This structure uses multibit feedback and is well suited for digital input. The most important design constraint for the modulator is the requirement

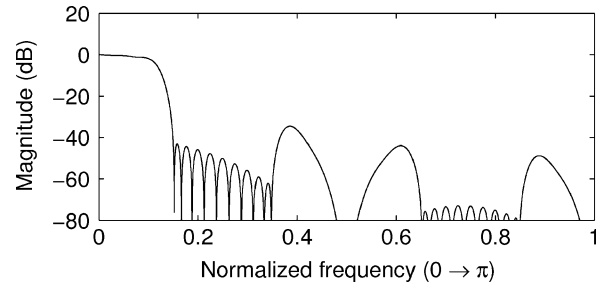


Fig. 5. Frequency response of the cascaded interpolation filter. Normalized to half the oversampling frequency.

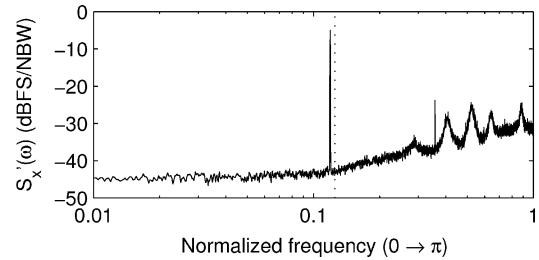


Fig. 6. Power spectral density of the sigma-delta modulator. Nyquist frequency indicated by dotted line.

that the output is a true bitstream. This rules out good modulator types such as cascaded or MASH architectures. These modulators combine several steps using filtering techniques following modulation. This requires at least one addition of several bitstreams, yielding word streams of more than one bit in width. Some experimentation was done using special bitstream adders with single-bit output. Common to these are that they either require an increased output rate (interleaving), or that they cause a slight deterioration of the signal quality, hence offsetting the benefit of a multistep modulator.

Quality of the modulator is also limited by the low OSR in the system. For an OSR of 8, modulator orders of two and more give a maximum expected Signal to Quantization Noise Ratio (SQNR) of 35–40 dB.

3) *Decimation Filter*: The filter used for decimation is an ordinary CIC decimation filter of third order with a downsampling ratio of 8. This converts the oversampled output from the cross-correlation block back to a Nyquist rate signal, while reducing high frequency noise inherited from the bitstream representation.

V. MEASUREMENT RESULTS

The chip is measured using a simple microcontroller, enabling measurements of the different chip modules.

1) *Interpolation Filter*: The frequency response of the cascaded interpolation filter is shown in Fig. 5. The response of the FIR filter step was confirmed by chip measurements, using white noise input signals and sinusoidal inputs. Total stop band attenuation is 35–40 dB, close to expected performance.

2) *Sigma-Delta Modulator*: The chip modulator was tested using a full-scale sinusoidal input. The output spectrum is shown in Fig. 6 and shows the expected SQNR level.

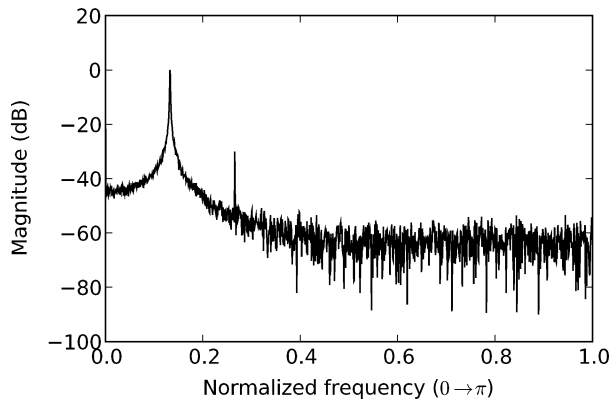


Fig. 7. Chip results showing the frequency spectrum of the bitstream cross-correlation between two sinusoids, decimated to the Nyquist rate.

3) *Bitstream Cross-Correlator*: To test the cross-correlator in isolation, a software sigma-delta modulator was used instead of the on-chip modulator. A full-scale sinusoidal input was modulated to a bitstream with SQNR ≈ 26 dB. The template register was loaded and the sinusoid was cross-correlated with a repeating version of itself. The result is downsampled, using a first order decimation filter. The frequency spectrum of the final Nyquist rate result is plotted in Fig. 7, showing a SNR ≈ 29 dB.

A. Beat Detection

A number of automated methods for measuring intervals and segments in the cardiac cycle have been developed over the years. The lack of standardized databases containing a sufficiently large number of manually annotated heartbeats has prevented good evaluation of the performance of these algorithms. Tremendous effort from clinicians is required to manually annotate a statistically significant set of ECG records. The QT Database (qtdb) [12] includes ECG to challenge beat detection algorithms with real-world variability. Cardiologists have manually annotated sections from each record in the database.

The multicomponent-based cross-correlation method for beat detection was benchmarked in [4], showing an improvement compared to other beat detection approaches. The detection of waveforms, performed by the low-power bitstream cross-correlator, is presented throughout this section.

1) *Template Creation*: Each template was created by averaging the ten first identifiable components of each type. Fig. 8 shows the QRS complex and T wave templates, accompanied with the data which they were generated from. The vertical lines show the boundaries of the template, in this measurement setup confined to a width of 64 samples or 256 ms. The templates have to be the correct length to optimally identify the target pattern. The fixed template length of 256 ms is not optimal, but close to the suggested correlation interval of 300 ms in [4].

2) *Beat Detection*: The first objective is to localize the QRS complex. The decimated chip result is plotted in Fig. 9 together with the “ideal” cross-correlation, where “ideal” is the mathematically correct cross-correlation between the ECG signal and the template. The “QRS” labels and the associated lines show the markers inserted by clinical experts. Solely by inspection, the bitstream cross-correlation results seem promising and very

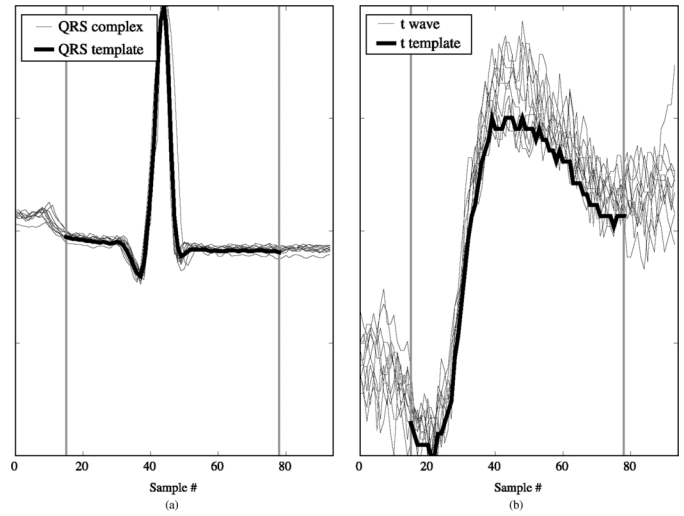


Fig. 8. Two of the used templates. (a) QRS wave template. (b) T wave template.

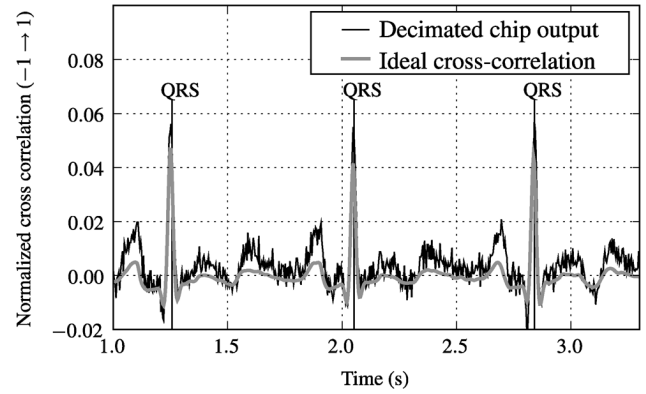


Fig. 9. Cross-correlation results for the QRS complex.

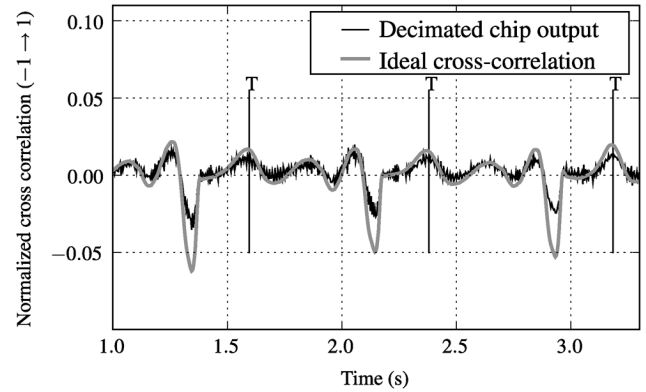


Fig. 10. Cross-correlation results for the T wave.

close to the “ideal” result. The T wave result, plotted in Fig. 10, shows that the template chosen for this quick test of the system did not enable proper detection of the T wave. The cross-correlation operation by itself, though, seems fine and with a proper algorithm, as that in [4], the T wave should be detectable well within the accepted detection tolerance.

The chip is evaluated for detection of the QRS complex in the 30 annotated heartbeats in the record *sel100* from the qtdb. The QRS peaks were localized within a standard deviation of

3.6 ms from the expert annotations. The reduction of accuracy compared to the standard deviation of 1.8 ms reported in [4] is mostly a result of the coarse delta-sigma modulator used for producing the bitstream, and not the bitstream operation itself. The result is noisier than ideal cross-correlation, but can be improved to some extent by using a modulator giving a higher signal-to-noise ratio (e.g., using a higher order modulator and a higher OSR) at the cost of higher power consumption. Sensor data are often sampled by a delta-sigma modulator, so bitstream operations can be used to reduce processing demands without any significant loss of signal quality. This first attempt at bitstream multicomponent-based beat detection seems promising. A proper evaluation of the bitstream cross-correlation method for detection of components in an ECG signal demands extensive analysis, and has not yet been performed.

B. System Performance

1) *Power Consumption:* The power consumption of the bitstream cross-correlator is proportional to the data processed. With an average load, the chip power consumption is measured to 5.1 mW, producing 120 k cross-correlation results each second at Nyquist rate. This is with $n_{\text{Nyquist}} = 128$, oversampling ratio = 8, and, consequently, $n_{\text{bitstream}} = 1024$.

The bitstream cross-correlator can be compared to the high performance, low power, 8-bit microcontroller (MCU) Atmel ATmega48PA with respect to power efficiency. The MCU was set to compute cross-correlation on binary coded data, with all possible power saving modes turned on. The MCU instruction set documentation shows that template loading, counter increment, multiply and accumulate and branching takes 10 clock cycles. With $n_{\text{Nyquist}} = 128$, this adds up to 1280 clock cycles per cross correlation result. The highest MCU clock frequency is 20 MHz, dissipating 57 mW, resulting in close to 16 k cross-correlations per second. The figure of merit used is $\text{FOM} = (\text{number of operations per second})/(\text{power dissipated})$. The improvement in power efficiency by using bitstream cross-correlation is then given by

$$\text{FOM}_{\text{BS}}/\text{FOM}_{\text{MCU}} = \frac{120 \text{ kOp/s}}{5.1 \text{ mW}} \bigg/ \frac{16 \text{ kOp/s}}{57 \text{ mW}} \approx 84$$

This rough comparison to a low power microcontroller shows a power consumption improvement of 84 times. However, more power-efficient MCUs have been introduced since this comparison was done. Also, digital signal processors might be more power-efficient than this prototype bitstream cross-correlator, which after all is just an exploration of new signal processing concepts and not an attempt at immediate power efficiency improvement. Bitstream operations, instead of processing of binary coded data, is a concept not fully explored and is a promising way of utilizing the internal data representation of a sigma-delta modulator. The analysis was performed under the assumption that a bitstream representation of the ECG signal was already available and without decimation of the cross-correlation result.

Simulated power dissipation results for the signal conversion blocks are summed up in Table I, along with measured results for the cross-correlator block. The results indicate that static

TABLE I
SIMULATED POWER DISSIPATION PER BLOCK $f_s = 18.8 \text{ kHz}$

Circuit	P_{static}	P_{total}
DSM	1.5 μW	1.5 μW
FIR filter	7.1 μW	7.1 μW
CIC decimator	3.7 μW	3.7 μW
CIC interpolator	2.5 μW	2.5 μW
Cross-correlator*	650 μW	730 μW
Total	665 μW	745 μW

*Measured values.

TABLE II
SUMMARY OF MAIN CHARACTERISTICS

Property	Value
Technology	90 nm CMOS
Area	1.0 mm ²
Supply voltage	1.0 V
Power consumption at 18.8 kHz	745 μW
Correlation window length	1024 bit
Converter SNR	25–30 dB
Max operating frequency	6.8 MHz

power dissipation is dominating when operating at low frequencies.

2) *Speed Performance:* Remember that the bitstream multiplication operation is done in parallel and that the multiplication result is sorted asynchronously. The sorting of one bit, including margins, is simulated to 144 ps and the worst case number of parallel sorting operations is given by the bubble register length -1 . This gives a theoretical maximum clock frequency of $f_{\text{max}} = 6.8 \text{ MHz}$.

A Nyquist rate cross-correlation result is ready every OSR clock periods, giving Nyquist rate cross-correlations per second = $f_{\text{max}}/\text{OSR}$. Substituting the system parameters, $f_{\text{max}} = 6.8 \text{ MHz}$ and $\text{OSR} = 8$, gives 850 k cross-correlation values each second. The achievable clock frequency is most likely not a limitation when processing real-time data, the register length confines the template length. The theoretical maximum speed restricts the template length to 0.15 ms. A more appropriate Nyquist rate could be 250 Hz, giving room for a template of more than 0.5 s. This sample rate is suitable when sampling heartbeats, which enables the implemented cross-correlator to be used for heartbeat detection.

VI. CONCLUSION

In this paper, we have presented a novel bitstream-based single-chip running cross-correlator implemented in 90 nm CMOS technology. The key system characteristics are summarized in Table II. The compact and power-efficient principles are intended for body sensor network applications or similar battery-operated, template-matching applications. The proposed cross-correlator is experimentally verified for component-based ECG analysis and a power improvement is observed compared to a generic MCU implementation. Dedicated ECG analysis hardware embedded in the ECG-electrode itself will reduce communication demands and power consumption in BSNs significantly.

Although we have used ECG-analysis as an example in this paper, the bitstream-based template-matching is an attractive

function usable in other sensor data processing applications. The cross-correlator may be used for convolution by reversing the template before loading the template register. Our solutions may also be well suited for low-frequency filtering operation, since sigma-delta noise-shaping facilitates low in-band quantization noise. Heart-rate variability analysis is one example of filtering studied in [13].

ACKNOWLEDGMENT

The authors would like to thank Dept. of Informatics, University of Oslo, for providing resources for chip fabrication and lab facilities for this project.

REFERENCES

- [1] G. Yang, *Body Sensor Networks*. New York: Springer, 2006.
- [2] B. Warneke, M. Last, B. Liebowitz, and K. S. J. Pister, "Smart dust: Communicating with a cubic-millimeter computer," *Computer*, vol. 34, no. 1, pp. 44–51, Jan. 2001.
- [3] B.-U. Kohler, C. Hennig, and R. Orglmeister, "The principles of software qrs detection," *Eng. Med. Biol. Mag.*, vol. 21, no. 1, pp. 42–57, Feb. 2002, IEEE.
- [4] T. Last, C. Nugent, and F. Owens, "Multi-component based cross correlation beat detection in electrocardiogram analysis," *BioMed. Eng. OnLine* vol. 3, no. 1, p. 26, 2004. [Online]. Available: <http://www.biomedical-engineering-online.com/content/3/1/26>
- [5] P. O'Leary and F. Maloberti, "Bit stream adder for oversampling coded data," *Electron. Lett.*, vol. 26, pp. 1708–1709, Sep. 1990.
- [6] F. Maloberti and P. O'Leary, "Processing of signals in their oversampled delta-sigma domain," in *Proc. Circuits and Systems, Int. Conf.*, China, Jun. 1991, vol. 1, pp. 438–441.
- [7] F. Maloberti, "Non conventional signal processing by the use of sigma delta technique: A tutorial introduction," in *Proc. IEEE Int. Symp. Circuits and System*, May 1992, vol. 6, pp. 2645–2648.
- [8] S. Summerfield, S. M. Kershaw, and M. B. Sandler, "Sigma-delta bit-stream filtering in vlsi," in *Proc. 37th Midwest Symp. Circuits and Systems*, Aug. 1994, vol. 2, pp. 1200–1203, vol. 2.
- [9] E. Janssen and D. Reefman, "Super-audio cd: An introduction," *IEEE Signal Proces. Mag.*, vol. 20, no. 4, pp. 83–90, Jul. 2003.
- [10] T. S. Lande, T. G. Constantinou, A. Burdett, and C. Toumazou, "Running cross-correlation using bitstream processing," *Electron. Lett.*, vol. 43, no. 22, 2007.
- [11] E. B. Hogenauer, "An economical class of digital filters for decimation and interpolation," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-29, no. 2, pp. 155–162, Apr. 1981.
- [12] P. Laguna, R. Mark, A. Goldberger, and G. Moody, "A database for the evaluation of algorithms for measurement of qt and other waveform intervals in the ecg," *Comput. Cardiol.*, vol. 24, pp. 673–676, 1997.
- [13] M. M. S. Ho, T. S. Lande, and C. Toumazou, "Efficient computation of the lf/hf ratio in heart rate variability analysis based on bitstream filtering," in *Proc. Biomedical Circuits Systems Conf.*, Nov. 2007, pp. 17–20.



Olav E. Liseth received the M.Inf. degree in microelectronics from the University of Oslo, Oslo, Norway, in 2009.

Currently, he is an IC Design Engineer with the university spin-off company Novelda, as developing clockless digital circuitry in state-of-the-art, low-power, high-resolution ultrawideband radar systems.



Daniel Mo received the M.Inf. degree in microelectronics from the University of Oslo, Oslo, Norway, in 2009.

Currently, he is an Electronics Design Engineer at Idex ASA, developing and testing integrated circuits for fingerprint sensors.



Håkon A. Hjortland (M'09) received the M.Inf. degree in microelectronics from the University of Oslo, Oslo, Norway, in 2006, where he is currently pursuing the Ph.D. degree.

His research interests include ultrawideband impulse radio, short-range radar, and wireless data communication.



in two books.

Prof. Lande is an IEEE Distinguished Lecturer and IEEE Fellow.

Tor Sverre "Bassen" Lande (M'93–SM'06–F'10) is a Professor of Microelectronics in the Department of Informatics, University of Oslo, and Visiting Professor at the Institute of Biomedical Engineering, Imperial College, London, U.K. His primary research is related to microelectronics, both digital and analog. His research interests include neuromorphic engineering, analog signal processing, micropower circuit design, biomedical circuits, as well as systems and impulse radio. He is the author or co-author of more than 100 scientific publications with chapters



Dag T. Wisland (M'96) received the M.Sc. and Dr.Scient. degrees in electrical engineering from the University of Oslo, Oslo, Norway, in 1996 and 2003, respectively.

Currently, he is CEO of the spin-off company Novelda AS and is Associate Professor in the Nanoelectronics Group at the University of Oslo. From 2004 to 2008, he headed the Nanoelectronics Research Group at the University of Oslo. His research interests include low-power analog/mixed-signal complementary metal–oxide semiconductor design, ultrawideband radio, and the design of analog-to-digital converters/digital-to-analog converters with a particular focus on delta-sigma data converters. In his research, he has focused on conceptually new methods and topologies combined with low-power design.

A.5 Pulse Generator, N-Path Filter, Quantizer, and Power Harvesting

Paper 21 A novel 6.5 pJ/pulse impulse radio pulse generator for RFID tags

K. K. Lee, M. Z. Dooghabadi, H. A. Hjortland, O. Naess, and T. S. Lande. “A novel 6.5 pJ/pulse impulse radio pulse generator for RFID tags”. In: *Circuits and Systems (APCCAS), 2010 IEEE Asia Pacific Conference on*, pp. 184–187, Kuala Lumpur, Dec. 2010. doi:10.1109/APCCAS.2010.5775001. © 2010 IEEE. Reprinted with permission.

My Contributions to This Paper

I contributed to the creation of the pulse generator principles used in this paper.

A Novel 6.5 pJ/Pulse Impulse Radio Pulse Generator for RFID Tags

Kin Keung Lee, Malihe Zarre Dooghabadi, Håkon A. Hjortland, Øivind Næss and Tor Sverre “Bassen” Lande
Department of Informatics, University of Oslo, N-0316 Oslo, Norway
E-mail: kkleee@ifi.uio.no

Abstract—A novel impulse radio (IR) ultra wideband (UWB) pulse generator (PG) intended for RFID tags is presented. A new pulse-shaping approach suited for CMOS implementation is proposed. The power consumption and chip area are reduced compared to the conventional higher order Gaussian PGs. The proposed PG uses digital gates for timing achieving good power efficiency and, meanwhile, acceptable spectral filling. The circuit is scalable both in bandwidth and center frequency. The PG is designed in a TSMC 90 nm CMOS technology. Post-layout simulations show a worst-case power consumption from a 1.2 V supply to be 6.5 pJ/pulse for a 100 MHz pulse repetition frequency (PRF). The chip area is 0.00079 mm^2 ($38.2 \text{ } \mu\text{m} \times 20.8 \text{ } \mu\text{m}$) for the PG core.

Index Terms—IR, UWB, RFID, CMOS, Pulse generator

I. INTRODUCTION

Viable communication solutions have been pursued actively after FCC released a large UWB spectral mask (i.e. 3.1–10.6 GHz) for unlicensed use. Most available UWB technology is based on multi-band (OFDM) solutions for proper UWB spectral filling. Today we still do not find widespread use of UWB solutions.

Another immature technology is RFID tags. Most RFID tags nowadays are using narrowband technology and powered by a battery (or inductive coupling for short-range communication). The battery increases the production cost and the tag size significantly. One noticeable exception is a single-chip RFID tag solution from Tagent [1] with on-chip energy-harvesting circuits and a 6.7 GHz UWB transmitter. The power-harvesting technology is using a 5.8 GHz carrier and limited to one meter distance indicating power-limitations in the tag.

As a first step towards green (no battery) RFID tags, we are exploring a simple and very power-efficient IR-UWB PG, which is often the most power-consuming component of the transmitter, suitable for implementation in standard digital technology (CMOS). Unlike narrowband counterparts, IR-UWB PGs do not require precise carrier generators and the wideband transmission does not require accurate inter-pulse timing.

In this paper, a new pulse-shaping approach which requires only digital gates is proposed. The proposed pulse generator uses only digital gates for timing achieving good power efficiency and acceptable spectral filling. The circuit is scalable both in bandwidth (BW_{-10dB}) and center frequency (f_{center}) and is well suited for nanometer CMOS technology.

The PG performs very well, especially in term of power consumption and chip area, and is competitive to other recently published UWB PGs [2]–[5]. Post-layout simulations show a worst-case power consumption to be 6.5 pJ/pulse for a 100 MHz PRF. In these extremely power-limited applications, leakage current must also be accounted for, hence the gate area should be kept small.

II. PULSE-SHAPE GENERATION

Higher order Gaussian pulse-shapes are widely used in IR-UWB PGs to minimize the sideband interference, however, it is hard to generate higher order Gaussian pulses in standard CMOS processes with simple circuit solutions. A piecewise Gaussian approximation (PGA) approach was proposed in [6] and a promising low-power (5.6 pJ/pulse) IR-UWB PG using a similar approach was reported in [2]. The weakness is the circuit complexity, the number of generator stages and output drivers are proportional to the order of the Gaussian function. Only one output driver is used during each phase, the un-used output drivers add output capacitance and, hence, increase the power consumption. Also, in order to adopt to international spectral regulations (band group 6, i.e. 7.7–8.7 GHz, is the only band available worldwide [7]), a high-order Gaussian derivative (15-th) is required, but not feasible due to increased parasitic capacitance.

Instead of generating higher order Gaussian waveforms, we are proposing to use several simple pulses to construct a first order Gaussian envelope. BW_{-10dB} is then determined by the envelope width and f_{center} is determined by the frequency of the generated pulses. The somehow crude generated pulses will not add significant sideband interference as long as the pulse frequency is significantly higher than the envelope frequency. The idea is depicted in Fig. 1. The amplitudes of 5 GHz triangular and sine waveforms are shaped to make a Gaussian envelope with $\alpha = 0.6 \text{ ns}$. The frequency spectra are shown in Fig. 1c.

Loose-triangular pulses are selected because they can be generated easily in CMOS processes by charging and discharging a parasitic capacitor. It is difficult to model loose-triangular waveforms in mathematical simulators, however, its frequency spectrum should be somehow in between the spectra shown in Fig. 1c. f_{center} may be limited to be 5 GHz in this paper because of the technology limitations. We are expecting it can be pushed to higher frequencies, hence better spectral utilization can be achieved with faster CMOS technology.

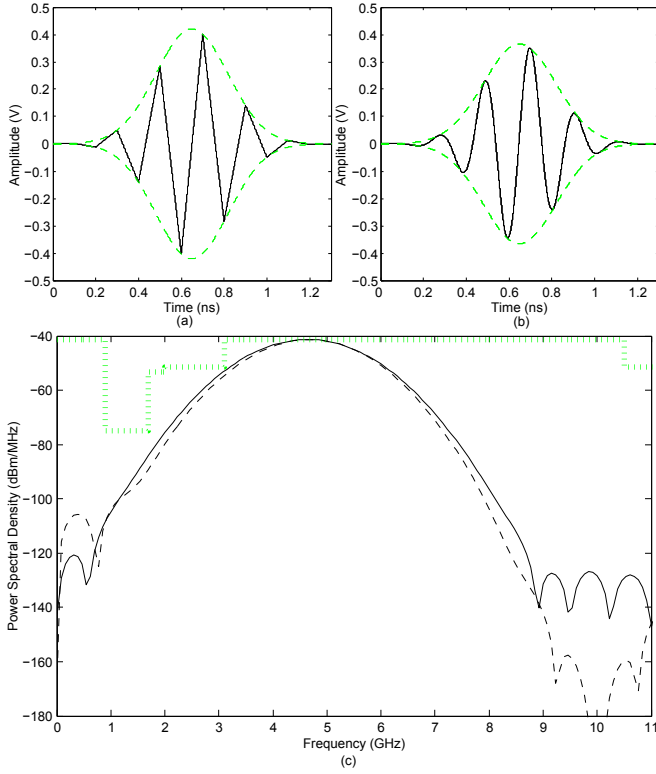


Fig. 1. (a and b) Proposed pulse-shape generation approaches (c) Dashed: Frequency spectrum of (a); Solid: Frequency spectrum of (b); Dotted: FCC mask.

The main advantage of this approach is its simplicity and technology scalability combined with power efficiency.

III. CIRCUIT IMPLEMENTATION

Because the gate count is proportional to the number of triangular pulses generated, the weak pulses at the beginning and at the end of the pulse-sequence are omitted in order to simplify the circuit and minimize the power consumption. The resulting waveform and spectrum is shown in Fig. 2. The f_{center} is 4.8 GHz with a BW_{-10dB} of 3 GHz. Some low-frequency sidebands are observed. However, simulations show that the mismatch between the charging and discharging time is the main reason for low-frequency sidebands and omission of the weak pulses will not have a significant impact on the sidebands. The low-frequency sidebands may be reduced by adding high-pass filters, or even just exploring the antenna band-pass property.

The schematic of the proposed PG is depicted in Fig. 3. The structure is similar to the PGA-PG in [2], but the idea is different. The proposed PG contains only simple digital gates and does not have any Gaussian pulse generators. The one-shot circuit is triggered by a falling edge and creates a fixed-width pulse, τ , for each input edge. Pulse stretchers [8] are added to adjust the width of the pulses in order to minimize the sideband energy. High threshold transistors are used for the delay line to minimize the leakage current.

As mentioned, in PGA-PGs, only one output driver is on

in a single phase, the un-used output drivers add undesired output capacitance and degrade the circuit performance. In the proposed PG, only loose-triangular waveforms are generated, this makes it possible to share the output drivers like the timing diagram shown in Fig. 3. One or more output drivers with different strength are combined for desired charging and discharging strength, which is a very important feature for generating multi-cycles IR-UWB pulses. The firing control logic is required for correct firing sequence. We can see that the PG contains only digital gates for timing exploring inherent gate-delays, which is very beneficial for modern CMOS technology.

IV. LAYOUT

The layout is shown in Fig. 4. Since the PG contains only digital gates, we use only 0.00079 mm² silicon area. The delay line is one of the critical components in the PG and must be designed carefully for minimal mismatch. Using multiple fingers devices and sharing the drain connections may reduce the parasitic drain-capacitance. However, this may give a very long diffusion and the “Length of Diffusion” effect [9] may degrade the delay line speed. Power distribution must also be balanced and decoupled locally since gates are rail referred. Just minor drops in power supply can affect inverter delays significantly.

V. POST-LAYOUT SIMULATION RESULTS

The PG testbench is shown in Fig. 5. The 100 fF and 10 μ F capacitors are the pad parasitic capacitor and coupling capacitor respectively. The 1 nH inductor accounts for the bond-wire parasitic inductance. The antenna is modeled as a 50 Ω resistive load. C_L is the loading capacitance due to the off-chip components (package, PCB and so on).

Fig. 6 shows the frequency spectra of the PG output (observed at V_{ANT}) with different C_L at the typical design corner. With 1 pF C_L , the peak is -41.9 dBm/MHz at around 4.3 GHz. The f_{center} is lower than the predicted because of the unexpected large parasitic capacitance of the delay line. The BW_{-10dB} is 2.9 GHz, which is very close to the predicted value. When C_L increases, the amplitude of triangular pulses decreases, which also decreases the peak of the spectra, the PG can meet the FCC requirements when

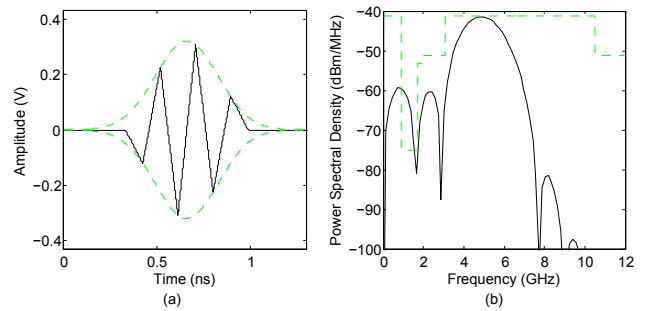


Fig. 2. (a) Adopted pulse-shape, (b) Solid: Its frequency spectrum; Dashed: FCC mask.

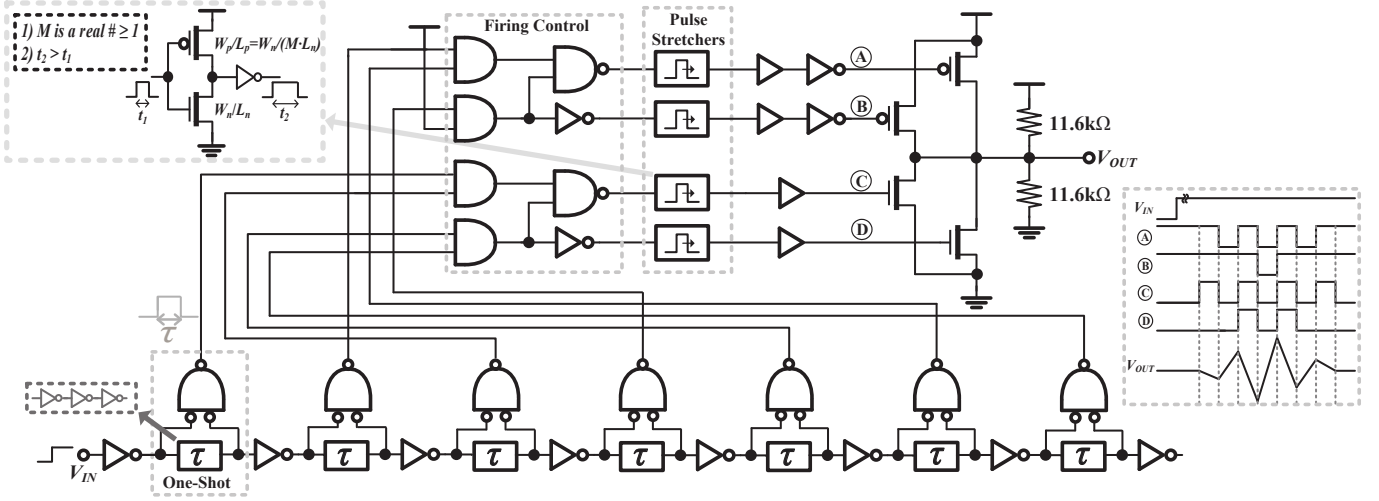


Fig. 3. Proposed PG.

C_L is larger than 1.5 pF assuming a high-pass filter added. With faster CMOS technology, we can push f_{center} to higher frequencies and achieve better spectral utilization

Fig. 7 shows a transient simulation with 2 pF C_L at the typical design corner. The pulse width (τ_{OUT}) is about 688 ps, and 820 ps for the worst case (i.e. the slow design corner). The ringing is caused by the bond-wire inductance. The

power consumption from a 1.2 V supply is 6.3 pJ/pulse for a 100 MHz PRF, and 6.5 pJ/pulse for the worst case. It can be seen in Fig. 7 that the maximum PRF can be larger than 200 MHz. The frequency spectra of the PG output (observed at V_{ANT}) with 2 pF C_L at different design corners are depicted in Fig. 8.

A summary of the simulation results and a comparison to recently published IR-UWB PGs [2]–[5] are shown in Table I. The simulations show that the designed PG performs very well, especially in terms of power dissipation and area, and is competitive to the state-of-the-art IR-UWB PGs.

VI. CONCLUSION

A new pulse-shaping approach and a novel IR-UWB PG for a TSMC 90 nm CMOS process suitable for RFID tags have been presented. A power-efficient pulse generation is achieved by output driver re-use and simple combinatorial gates. Post-layout simulations show the worst-case power consumption from a 1.2 V supply is 6.5 pJ/pulse for a 100 MHz PRF, which is suitable for ultra low-power applications. The PG contains

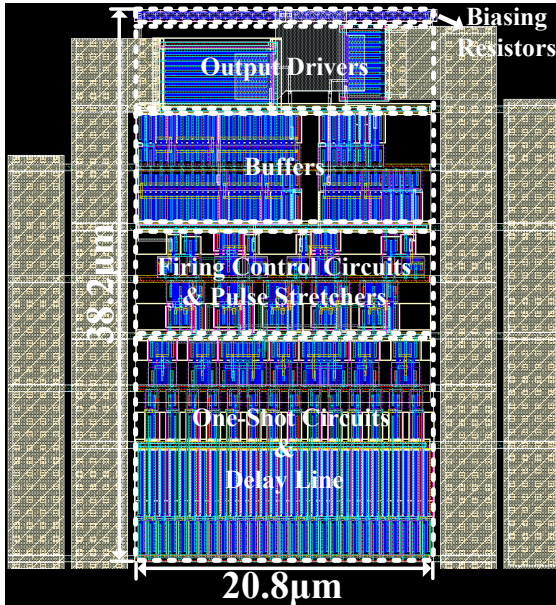


Fig. 4. PG layout.

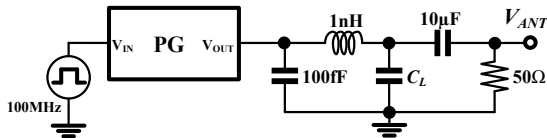


Fig. 5. PG testbench.

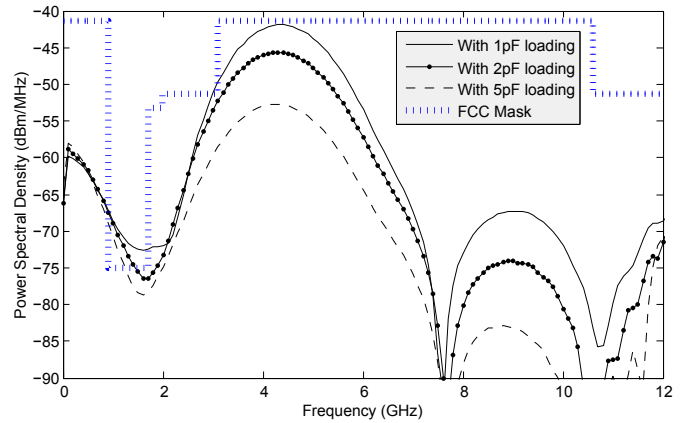
Fig. 6. Frequency spectra of the PG output with different C_L .

TABLE I
SUMMARY OF SIMULATION RESULTS, AND COMPARISONS TO OTHER PUBLISHED
UWB-IR PGs.

Design	Tech. (CMOS)	f_{center} (GHz)	BW_{-10dB} (GHz)	τ_{OUT} (ns)	Power Diss. (pJ/Pulse)	Area (mm ²)
This work	90 nm	4.3	2.94	0.69	6.5	0.00079
[2]	130 nm	~3.1	N/A	0.75	5.6	0.02
[3]	90 nm	4.05*	0.55	3	47	0.08
[4]	180 nm	4.05	1.4	1.75	825	0.4
[5]	90 nm	3.5–10	0.5	17	40	0.066

* Two other f_{center} (i.e. 3.45 GHz and 4.65 GHz) can be selected.

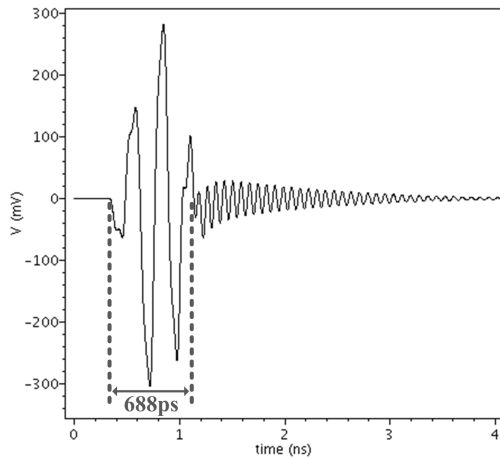


Fig. 7. Transient simulation result

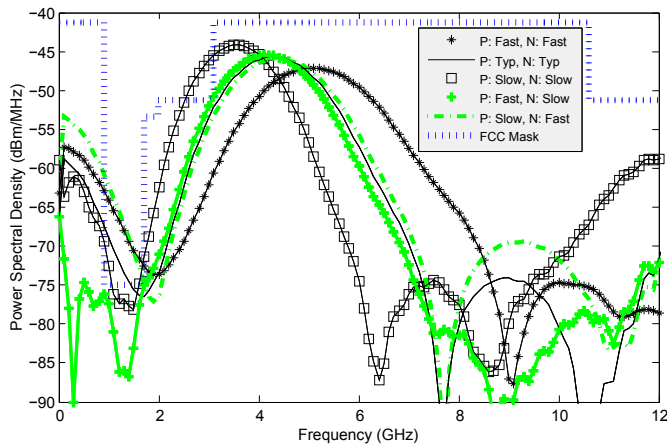


Fig. 8. Frequency spectra of the PG output at different design corners.

mainly digital gates, which is very beneficial from nano-meter CMOS technology. We are expecting this IR-UWB PG to be suitable for RFID tags with on-chip power-harvesting.

ACKNOWLEDGMENT

The authors would like to thank Novelda AS for their useful suggestions on IR-UWB PG design.

REFERENCES

[1] [Online]. Available: <http://www.tagent.com/>

- [2] X. Wang *et al.*, "FCC-EIRP-aware UWB pulse generator design approach," in *IEEE International Conference on Ultra-Wideband*, Sep 2009, pp. 592–596.
- [3] D. D. Wentzloff and A. P. Chandrakasan, "A 47pJ/pulse 3.1-to-5GHz all-digital UWB transmitter in 90nm CMOS," in *Digest of Technical Papers, IEEE International Solid-State Circuits Conference*, Feb 2007, pp. 118–119 and 591.
- [4] T. Norimatsu *et al.*, "A UWB-IR transmitter with digitally controlled pulse generator," *IEEE Journal of Solid-State Circuits*, vol. 42, no. 6, pp. 1300–1309, Jun 2007.
- [5] J. Ryckaert *et al.*, "A 0.65-to-1.4nJ/burst 3-to-10 GHz UWB all-digital TX in 90nm CMOS for IEEE 802.15.4a," *IEEE Journal of Solid-State Circuits*, vol. 42, no. 12, pp. 2860–2869, Dec 2007.
- [6] H. Hie *et al.*, "A varying pulse width 5th-derivative Gaussian pulse generator for UWB transceivers in CMOS," in *IEEE Radio and Wireless Symposium*, Jan 2008, pp. 171–174.
- [7] WiMedia Alliance, "Worldwide regulatory status – Jan 2009."
- [8] H. A. Hjortland and T. S. Lande, "CTBV integrated impulse radio design for biomedical applications," *IEEE Transactions on Biomedical Circuits and Systems*, vol. 3, no. 2, pp. 79–88, Apr 2009.
- [9] P. G. Drennan, M. L. Kniffin, and D. R. Locascio, "Implications of proximity effects for analog design," in *IEEE Custom Integrated Circuits Conference*, Sep 2006, pp. 169–176.

Paper 22 A 5.2 pJ/pulse impulse radio pulse generator in 90 nm CMOS

K. K. Lee, M. Z. Dooghabadi, H. A. Hjortland, O. Næss, and T. S. Lande. “A 5.2 pJ/pulse impulse radio pulse generator in 90 nm CMOS”. In: *Circuits and Systems (ISCAS), 2011 IEEE International Symposium on*, pp. 1299–1302, Rio de Janeiro, May 2011. doi:10.1109/ISCAS.2011.5937809. © 2011 IEEE. Reprinted with permission.

My Contributions to This Paper

I contributed to the creation of the pulse generator principles used in this paper, but not the dynamic pre-charge, though.

A 5.2 pJ/Pulse Impulse Radio Pulse Generator in 90 nm CMOS

Kin Keung Lee, Malihe Zarre Dooghabadi, Håkon A. Hjortland, Øivind Næss and Tor Sverre “Bassen” Lande
Department of Informatics, University of Oslo, N-0316 Oslo, Norway
E-mail: kkleee@ifi.uio.no

Abstract—A low-power impulse radio (IR) ultra wideband (UWB) pulse generator (PG) is presented. It uses digital gate-delay for timing achieving good power efficiency and acceptable spectral filling. The circuit is scalable in both bandwidth and center frequency. Both energy consumption and chip area are reduced compared to most conventional higher order Gaussian PGs. The PG is realized in a TSMC 90 nm CMOS process. Measurements show the energy consumption from a 1.2 V to be 5.2 pJ/pulse for a 200 MHz pulse repetition frequency (PRF). The core area is 0.0015 mm² (38 μm × 40 μm). Lastly, a dynamic pre-charge (DPC) scheme is proposed to eliminate the stand-by current and make the PG favourable for low-data-rate applications.

Index Terms—IR, UWB, RFID, CMOS, pulse generator

I. INTRODUCTION

Since Federal Communication Commission (FCC) released a large spectral mask (i.e. 3.1–10.6 GHz) for unlicensed use [1], UWB technology has been an attractive field of research. The huge spectral mask not only gives the benefit of higher data rate, but also reduced power consumption of transmitting circuits by using IR technology. One potential application is RFID tags. Conventional passive RFID tags utilize load or backscatter modulations which limits the number of channels. By using power-efficient IR-UWB technology, it is possible to apply different multiple access schemes (for example CDMA) on no-battery, energy-harvesting RFID tags. One of the most challenging tasks is to design low-power IR-UWB PG which is usually the most power-consuming component in the UWB transmitters.

A novel pulse-shape generation approach was proposed in [2], but no real prototype and measurement results were presented. In this paper, a low-power IR-UWB PG based on [2] was realized in a TSMC 90 nm CMOS process. It uses digital gate-delay for timing, achieving good power efficiency and acceptable spectral filling. The circuit is scalable in both bandwidth and center frequency. Both energy consumption and chip area are reduced compared to most conventional higher order Gaussian PGs. Measurements show the energy consumption from a 1.2 V supply to be 5.2 pJ/pulse for a 200 MHz PRF. The core area is 0.0015 mm² (38 μm × 40 μm). The PG performs very well, especially in terms of energy consumption and chip area, and is competitive to other recently published UWB PGs [3]–[6]. One drawback is that the output DC voltage is biased by a resistive divider. A DPC scheme is proposed to eliminate the stand-by current due to the

resistive divider and make the PG favorable for low-data-rate applications.

II. PULSE-SHAPE GENERATION

Higher order Gaussian pulse-shapes are widely used in IR-UWB PGs to minimize the sideband interference, however, it is hard to generate higher order Gaussian pulses in standard CMOS processes with simple circuit solutions. A piecewise Gaussian approximation (PGA) approach was proposed in [7] and a very low-power (5.6 pJ/pulse) IR-UWB PG using a

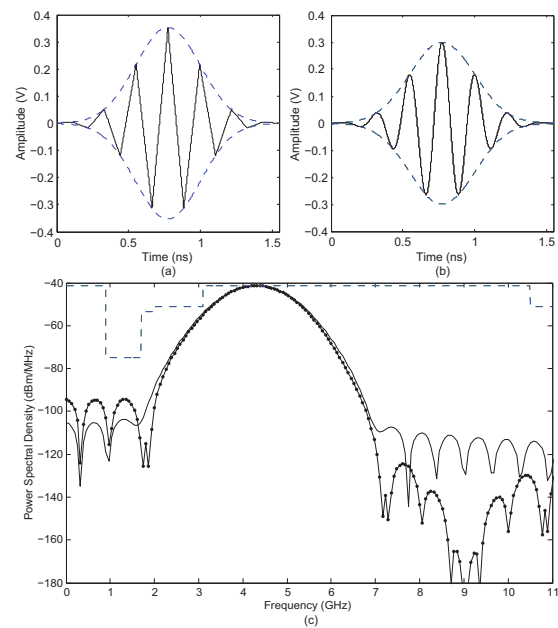


Fig. 1. (a & b) Pulse-shape generation approach. (c) Dotted: PSD of (a); Solid: PSD of (b); Dashed: FCC mask.

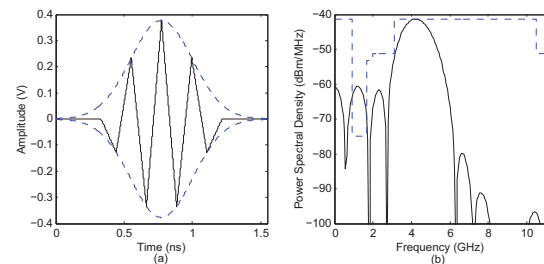


Fig. 2. (a) Adopted pulse-shape. (b) Solid: Its PSD; Dashed: FCC mask.

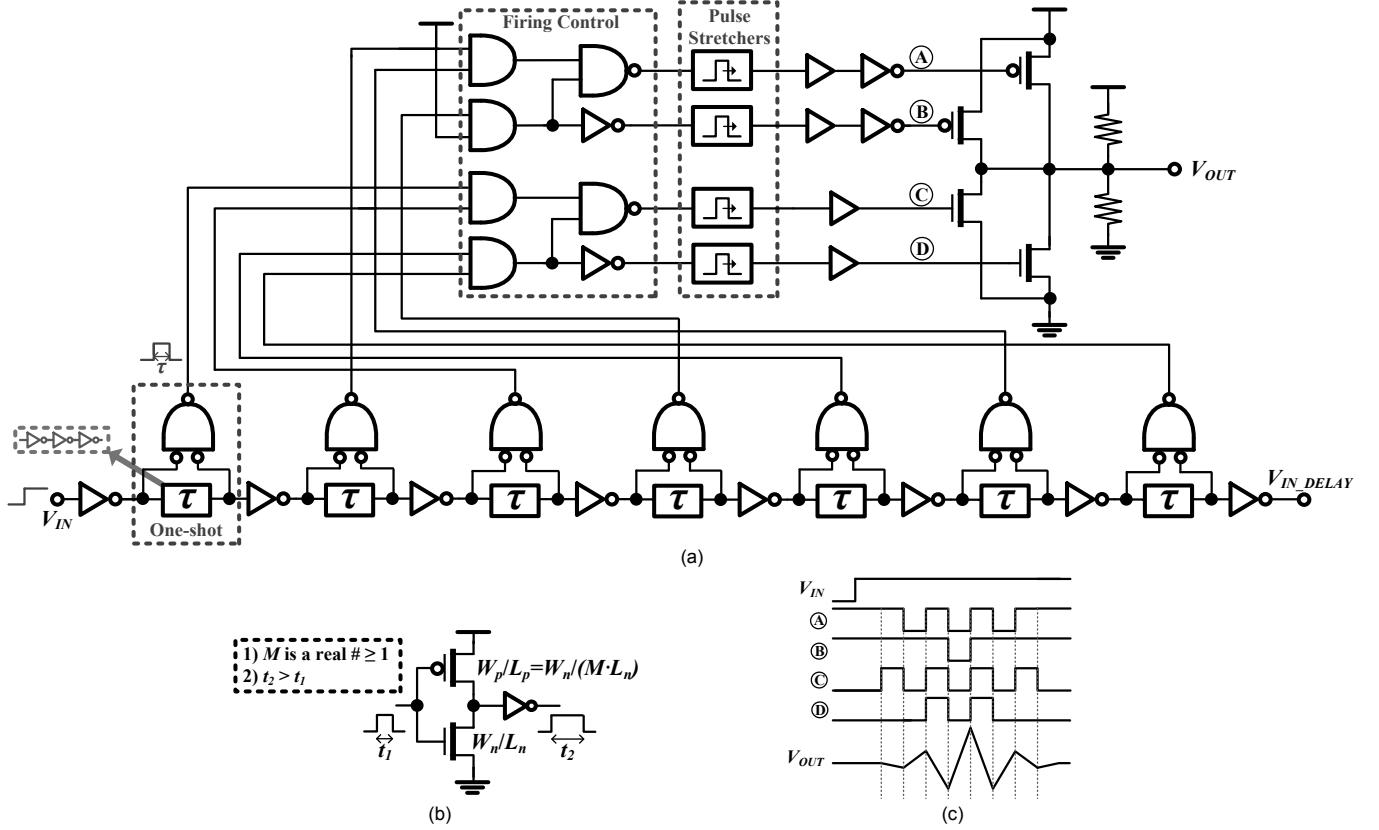


Fig. 3. (a) Designed PG. (b) Pulse stretcher. (c) Timing diagram.

similar approach was reported in [3]. The weakness is the circuit complexity, the number of generator stages and output drivers are proportional to the order of the Gaussian function. Only one output driver is used during each phase, the unused output drivers add output capacitance and, hence, increase the power consumption. Also, in order to adopt international spectral regulations (band group 6, i.e. 7.7–8.7 GHz, is the only band available worldwide [8]), a high-order Gaussian derivative (15-th) is required, but not feasible due to the increased parasitic capacitance.

Instead of generating higher order Gaussian waveforms, we use several simple pulses to construct a first order Gaussian envelope. The bandwidth (BW_{-10dB}) is then determined by the envelope width and the center frequency (f_{center}) is determined by the frequency of the generated pulses. The somehow crude generated pulses do not add significant sideband interference as long as the pulse frequency is significantly higher than the envelope frequency. The idea is depicted in Fig. 1. The amplitudes of 4.3 GHz triangular and sine waveforms are shaped to make a Gaussian envelope with $\alpha = 0.8$ ns. The power spectrum densities (PSD) are shown in Fig. 1c.

Loose-triangular pulses are selected in this paper because they can be generated easily in CMOS processes by charging and discharging a parasitic capacitor. It is difficult to model loose triangular waveforms in mathematical simulators, how-

ever, its PSD should be somehow in between those shown in Fig. 1c. f_{center} may be limited to be 4–5 GHz in this paper because of the technology limitations. We are expecting it can be pushed to higher frequencies, hence better spectral utilization can be achieved with faster CMOS technology. The main advantage of this approach is its simplicity and technology scalability combined with power efficiency.

III. CIRCUIT IMPLEMENTATION

Because the gate count is proportional to the number of triangular pulses generated, the weak pulses at the beginning and at the end of the pulse-sequence are omitted in order to simplify the circuit and minimize the power consumption. The resulting waveform and PSD is shown in Fig. 2. The f_{center} is 4.2 GHz with a BW_{-10dB} of 2.3 GHz. Some low-frequency sidebands are observed. However, simulations show that the mismatch between the charging and discharging time is the main reason for low-frequency sidebands and omission of the weak pulses does not have a significant impact on the sidebands. The low-frequency sidebands may be reduced by adding high-pass filters, or even just exploring the antenna band-pass property.

The schematic of the PG is depicted in Fig. 3a. The structure is similar to the PGA-PG in [3], but the idea is different. The designed PG contains only simple digital gates and does not have any Gaussian pulse generators. The one-shot circuit is

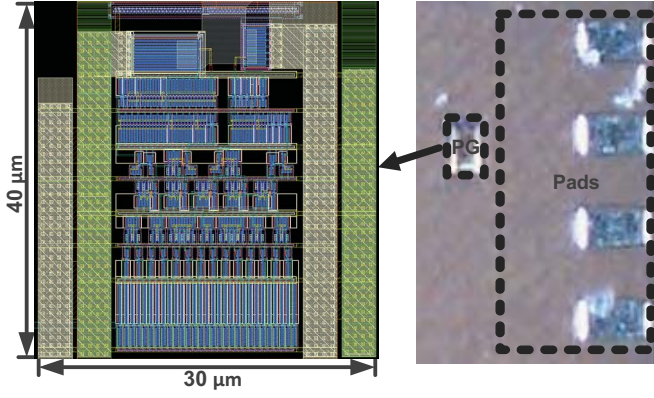
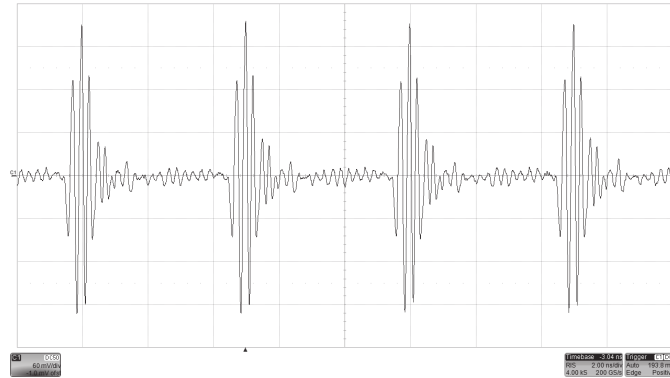


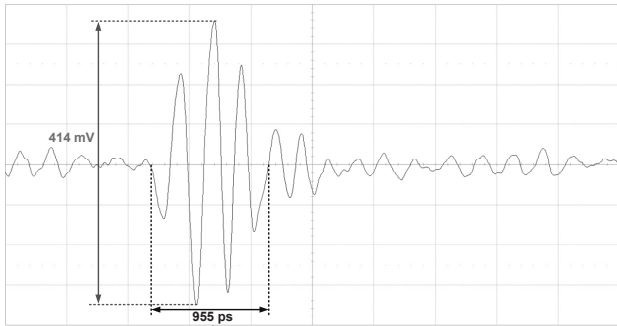
Fig. 4. (Left) Circuit Layout. (Right) Die photo.

triggered by a falling edge and creates a fixed-width pulse τ . Pulse stretchers [9] (schematic shown in Fig. 3b) are added to adjust the width of the pulses in order to minimize the sideband energy. High threshold transistors are used for the delay line to minimize the leakage current.

As mentioned, in PGA-PGs, only one output driver is on in a single phase, the unused output drivers add undesired output capacitance and degrade the circuit performance. In the designed PG, only loose-triangular waveforms are generated, this makes it possible to share the output drivers like the timing diagram shown in Fig. 3c. One or more output drivers



(a)



(b)

Fig. 5. (a) Output waveform for a 200 MHz PRF and (b) its zoom-in.

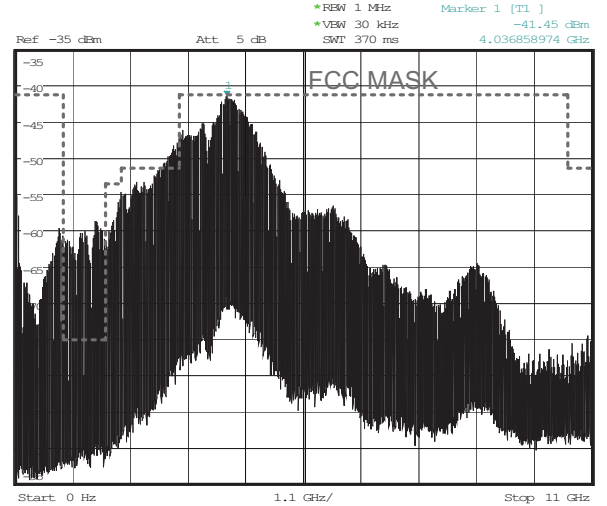


Fig. 6. Measured PSD.

with different strength are combined for desired charging and discharging strength, which is a very important feature for generating multi-cycles IR-UWB pulses. The firing control logic is required for correct firing sequence. The last pull-up is due to the output resistive divider. We can see that the PG contains only digital gates for timing, exploring inherent gate-delays.

IV. EXPERIMENTAL RESULTS

The PG is realized in a TSMC 90 nm CMOS process. A die photo is shown in Fig. 4. Since the PG contains mainly digital logics, we achieve a small silicon area of 0.0015 mm². The die is packaged in a 48-leads QFN package.

Fig. 5 shows the output waveforms for a 200 MHz PRF. The

TABLE I
MEASURED ENERGY CONSUMPTION AT DIFFERENT PRF.

PRF (MHz)	Energy Cons. (pJ/Pulse)
1	67.5
5	17.3
10	11.2
20	8.1
40	6.5
100	5.6
150	5.4
200	5.2

TABLE II
SIMULATED ENERGY CONSUMPTION WITH/WITHOUT DPC SCHEME
AT DIFFERENT PRF.

PRF (MHz)	Energy Cons. (pJ/Pulse)	
	w/ DPC	w/o DPC
0.5	5.4	129.1
1	5.3	66.9
5	5.2	17.2
10	5.2	11.0
20	5.2	7.9
50	5.1	6.0

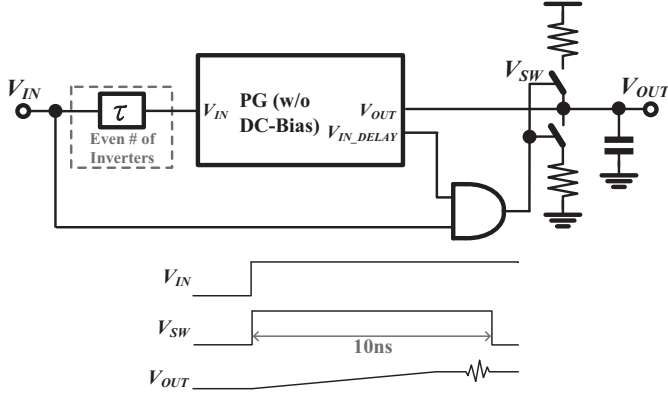


Fig. 7. DPC scheme.

pulse width (τ_{out}) is 955 ps and the peak-to-peak voltage is 414 mV, the ringing is due to the bondwire inductance.

Its PSD is depicted in Fig. 6. f_{center} is 4 GHz approximately and BW_{-10dB} is around 1.9 GHz (i.e. 2.89–4.81 GHz), which are very close to our expected values. With faster CMOS processes, we can push f_{center} to higher frequencies and achieve better spectral utilization.

Table I shows the measured energy consumption at different PRF. When the PRF is 200 MHz, the PG shows the lowest energy consumption of 5.2 pJ/pulse for a 1.2 V supply. However, the energy consumption increases when the PRF decreases. It is because the resistive divider is used to bias the output DC voltage and drives static current. This makes it difficult to apply the designed PG on low-data-rate applications.

V. DYNAMIC PRE-CHARGE

To eliminate the stand-by current consumption, a DPC scheme is proposed. The proposed circuit is shown in Fig. 7, the biasing resistors are connected to the output a short time before the pulse is transmitted and disconnected immediately after the radiation. The resistors and, hence, the time constant of pre-charging should be set reasonably large, otherwise the antenna may radiate unwanted EM energy during pre-charging.

Table II shows the pre-layout simulation results on the PG energy consumption with/without DPC. We can see that the stand-by energy consumption is greatly reduced with the DPC scheme and this makes the PG suitable for low-data-rate applications. The only stand-by component is now the leakage current. Thanks to the new efficient pulse-shape generation approach, we use only a very small silicon area and the leakage current is insignificant¹ compared to the total energy consumption.

VI. CONCLUSION

A low-power IR-UWB PG was presented. It was realized in a TSMC 90 nm CMOS process. It made use of a new pulse-shape generation algorithm which was suitable for CMOS im-

TABLE III

SUMMARY OF MEASUREMENT RESULTS, AND COMPARISON TO OTHER PUBLISHED IR-UWB PGs.

Design	Tech. (CMOS)	f_{center} (GHz)	BW_{-10dB} (GHz)	τ_{out} (ns)	Energy Cons. (pJ/Pulse)	Area (mm ²)
This work	90 nm	4	1.92	0.96	5.2	0.0015
[3]	130 nm	~3.1	N/A	0.75	5.6	0.02
[4]	90 nm	4.05*	0.55	3	47	0.08
[5]	130 nm	N/A	6.8	0.46	38.4	0.54
[6]	180 nm	~8	3.9	0.6	14	0.11

*Two other f_{center} (i.e. 3.45 GHz and 4.65 GHz) can be selected.

plementation. Measurements confirmed the energy consumption from a 1.2 V supply to be 5.2 pJ/pulse for a 200 MHz PRF. The core area was 0.0015 mm² (38 μ m \times 40 μ m). Also, a DPC scheme was proposed to eliminate the stand-by current and make the PG suitable for low-data-rate applications. A summary of the measurement results and comparison to some recently published IR-UWB PGs [3]–[6] are given in Table III. We can see that the PG performs very well, especially in term of energy consumption and chip area, and is competitive to other state-of-the-art IR-UWB PGs.

ACKNOWLEDGMENT

The authors would like to thank Novelda AS for their useful suggestions on IR-UWB PG design.

REFERENCES

- [1] Federal Communication Commission, *Revision of Part 15 of the Commission's Rules Regarding Ultra-Wideband Transmission Systems*, adopted Feb. 2002, released Apr. 2002.
- [2] K. K. Lee, M. Z. Dooghabadi, H. A. Hjortland, Ø. Næss, and T. S. Lande, "A novel 6.5 pJ/pulse impulse radio pulse generator for RFID tags," in *Proc. Asia Pacific Conference on Circuits and Systems*, Dec 2010, pp. 184–187.
- [3] X. Wang *et al.*, "FCC-EIRP-aware UWB pulse generator design approach," in *Proc. IEEE International Conference on Ultra-Wideband*, Sep 2009, pp. 592–596.
- [4] D. D. Wentzloff and A. P. Chandrakasan, "A 47pJ/pulse 3.1-to-5GHz all-digital UWB transmitter in 90nm CMOS," in *Digest of Technical Papers, IEEE International Solid-State Circuits Conference*, Feb 2007, pp. 118–119 & 591.
- [5] S. Bourdel *et al.*, "A 9-pJ/pulse 1.42-Vpp OOK CMOS UWB pulse generator for the 3.1–10.6-GHz FCC band," *IEEE Transactions on Microwave Theory and Techniques*, vol. 58, no. 1, pp. 65–73, Jan 2010.
- [6] S. Sim, D.-W. Kim, and S. Hong, "A CMOS UWB pulse generator for 6–10 GHz applications," *IEEE Microwave and Wireless Components Letters*, vol. 19, no. 22, pp. 83–85, Feb 2009.
- [7] H. Hie *et al.*, "A varying pulse width 5th-derivative Gaussian pulse generator for UWB transceivers in CMOS," in *Proc. IEEE Radio and Wireless Symposium*, Jan 2008, pp. 171–174.
- [8] WiMedia Alliance, "Worldwide regulatory status – Jan 2009."
- [9] H. A. Hjortland and T. S. Lande, "CTBV integrated impulse radio design for biomedical applications," *IEEE Transactions on Biomedical Circuits and Systems*, vol. 3, no. 2, pp. 79–88, Apr 2009.

¹The leakage current is usually proportional to the transistor width. For digital circuits, minimum length is commonly used, which means the total leakage current is proportional to the silicon area.

Paper 23 An inductorless 6-Path band-pass filter with tunable center frequency for UWB applications

T. A. Vu, S. Sudalaiyandi, H. A. Hjortland, O. Nass, T. S. Lande, and S. E. Hamran. “An inductorless 6-Path band-pass filter with tunable center frequency for UWB applications”. In: *Ultra-Wideband (ICUWB), 2012 IEEE International Conference on*, pp. 164–167, Syracuse, NY, Sep. 2012. doi:10.1109/ICUWB.2012.6340442. © 2012 IEEE. Reprinted with permission.

My Contributions to This Paper

This paper presents an N-path filter for microwave frequencies, suitable for UWB applications. The filter uses a 6-phase clock generator which utilizes both rising and falling edges in a three-inverter ring oscillator for increased operating frequency. I came up with the design for this clock generator. I also helped review the paper.

An Inductorless 6-Path Band-Pass Filter with Tunable Center Frequency for UWB Applications

Tuan Anh Vu, Shanthi Sudalaiyandi, Håkon A. Hjortland,
Øivind Næss, Tor Sverre Lande, and Svein Erik Hamran
Dept. of Informatics, University of Oslo, Norway
Email: anhtv@ifi.uio.no

Abstract—This paper presents an inductorless switched-capacitor 6-path band-pass filter suitable for impulse radio ultra-wideband (IR-UWB) application designed in nanometer CMOS technology. The proposed wideband band-pass filter is working in discrete-time intended for power-efficient IR-UWB systems. In 90 nm CMOS we were able to design a filter with -3 dB bandwidth of 2.5 GHz and a center frequency tunable from 4.1 GHz to 4.5 GHz. Estimated power consumption is 35 mW from 1.2 V supply voltage according to post-layout simulation. A novel ring oscillator based multiphase clock is proposed, enabling center-frequency tuning at microwave frequencies.

Index Terms—Band-Pass Filter, Center-Frequency Tuning, Multiphase Clock, N-Path Filter, Switched-Capacitor, UWB Filter.

I. INTRODUCTION

As wireless communication is moving up to higher frequencies, suitable integrated band-pass filters are required. Typically passive filters like Surface Acoustic Wave (SAW) filters are used. Aiming at integrated filter solutions, SAW-filters are not viable [1].

An additional challenge is wideband, microwave filters for impulse radio applications. Wideband resonator based filters are hard to make and area consuming. Recently, there has been a renewed interest in switched-capacitor (SC) N-path filters due to the development of CMOS technology allowing this type of filter to work at RF frequencies [2][3]. SC filters provide a practical method for fully integrated realization of high quality RF filters. N-path filters were originally designed for continuous-time band-pass filtering [3][4]. A combination of SC filter and N-path may be used as tunable, integrated filters in standard CMOS technology.

II. SWITCHED-CAPACITOR N-PATH FILTERS

A. Principles of SC N-Path Filters

A typical N-path filter is composed of N parallel identical filter cells which are cyclically sampled with the frequency F_s . Each path usually in turn consists of a passive low-pass filter (LPF) with transfer function $H(f)$ between an input and output mixer as shown in Fig. 1. The switches are driven by time/phase shifted versions of clock $p(t)$ and $q(t)$. When the switching signals are assumed to be ideal Dirac impulses delayed by T_s/N , the output signal $S(f)$ may be written as a function of the input signal $E(f)$ as shown in equation (1).

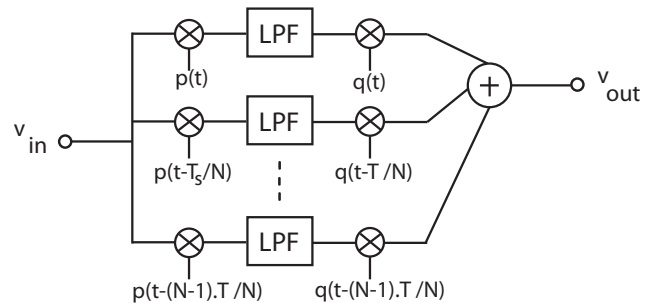


Fig. 1. The structure of an N-path filter.

$$S(f) = \sum_{m=-\infty}^{+\infty} h(m \cdot T_s) \cdot \exp(-2 \cdot \pi \cdot j \cdot f \cdot m \cdot T_s) \quad (1)$$

$$\cdot \sum_{n=-\infty}^{+\infty} e\left(\frac{n \cdot T_s}{N}\right) \cdot \exp(-2 \cdot \pi \cdot j \cdot f \cdot \frac{n \cdot T_s}{N})$$

Equation (1) shows that the output signal is the result of two multiplied sums. The first term corresponds to the Fourier transform of the impulse response of the low-pass filter sampled at T_s period. The second one corresponds to the Fourier transform of the input signal sampled at T_s/N period. Due to the periodicity of the frequency response, the band-pass center frequency is equal to a multiple of the clock frequency. This filter can be used for clock recovery by filtering the harmonic components [5]. It can also be used as band-pass filter centered around F_s [1]. Thus, the N-path filter transfers a low-pass characteristic to a band-pass whose center frequency is controlled by the clock frequency while the low-pass filter and duty cycle of the clock define the bandwidth. In fact the input signal is downconverted to baseband, filtered and then upconverted again to the same band as the input signal. Due to the Nyquist theorem, sampled-data filters can only process input signals up to half of the clock frequency. Since the transfer characteristic is the same for every switch position, the overall frequency response is equal to the response of an individual path. However, the output signal is now composed of N samples per period T_s , so that the Nyquist range NY for the N-path filter is expanded N times as follow:

$$NY = N \cdot \frac{1}{2} \cdot F_s \quad (2)$$

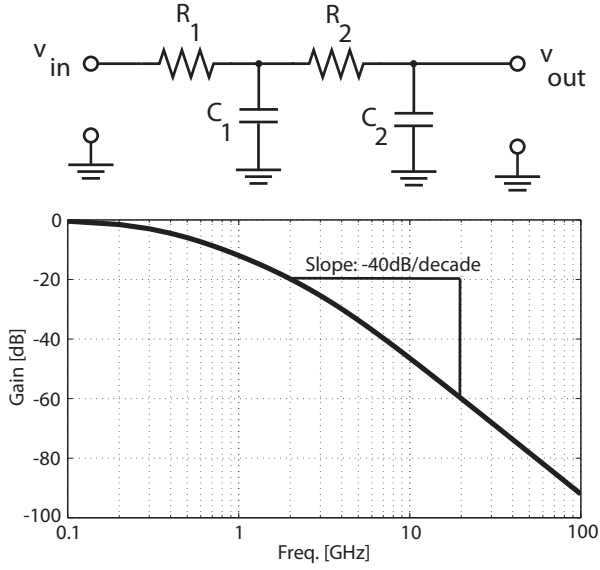


Fig. 2. A second-order passive low-pass filter and its frequency response.

The decomposition of filter center frequency and the filter properties is promising. A multi-phase clock generator with wide tuning-range is required for center frequency tuning. The filter sharpness is determined by the sharpness of the low-pass filter. For most wideband filters a second-order passive low-pass filter is sufficient.

A dominating error-source is clock feedthrough due to the parasitic capacitors. This clock feedthrough distortion produces not only an undesirable offset voltage but also additional spectral components at harmonics of the clock frequency F_s . These spectral components reduce the maximum dynamic range in the filter passband at F_s . Even harmonics may be eliminated by using differential signalling [3]. The clock feedthrough is therefore a major challenge designing SC N-path band-pass filters. In addition, due to tolerances in the filter paths, unwanted mirror frequencies can appear at the output [6].

B. Low-Pass Filter and Multiphase Clock Generator

As mentioned before, the order and the corner frequency of the low-pass filter defines the roll-off and bandwidth of the N-path band-pass filter. A second-order RC low-pass filter is designed for improved frequency roll-off as shown in Fig. 2. Thus, the gain decreases by 40 dB per decade above cut-off frequency. The frequency response of the second-order passive low-pass filter is given in equation (3).

$$\frac{v_{out}}{v_{in}} = \frac{\frac{1}{R_1 R_2 C_1 C_2}}{s^2 + s \left(\frac{1}{R_1 C_1} + \frac{1}{R_2 C_1} + \frac{1}{R_2 C_2} \right) + \frac{1}{R_1 R_2 C_1 C_2}} \quad (3)$$

Since the center frequency of the SC N-path band-pass filter can be tuned with the clock frequency, a tunable oscillator is desirable. The tunability is particularly useful for RF integrated circuits, for which the tuning of the center frequency

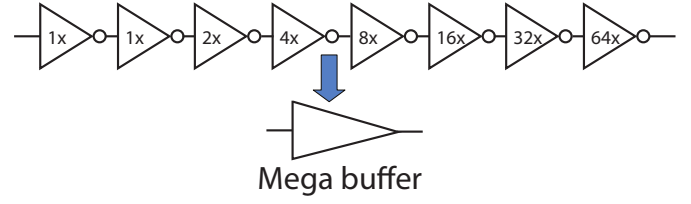


Fig. 3. The buffer with suitable fan-out.

by a clock signal permits either to compensate for process variations and device mismatch or to filter various channels with the same filter. Several clock generator topologies for RF frequencies have been proposed in literature [7][8]. Most of them use inductors and capacitors to provide oscillation. These inductors occupy a very large area in the layout and are hard to adapt to finer pitch technology.

In this paper, a novel ring oscillator based, multiphase clock generator is proposed. Avoiding an external crystal-driven clock, higher frequencies are feasible and full integration is possible. The multiphase clock generator is built with digital logic gates and can be scaled with CMOS technology. A minimum three-stage ring oscillator provides oscillation resulting in a very high-frequency clock suitable for UWB applications. Since the frequency of oscillation depends on the inherent inverter delay plus whatever additional loading is applied, taps from the delay line must have minimal load for maximum operation frequency. Consequently, a buffer with very low input capacitance is required to minimize the load capacitance while having the capability of driving large loads at the output. Fig. 3 shows a buffer with incremental loading for optimal performance.

Fundamentally, the number of inverters in the ring oscillator equals to the number of clock phases needed. However, both rising and falling edges present in the ring oscillator can be exploited for twice as many clock phases. The multiphase clock generator is shown in Fig. 4. In order to tune the oscillation frequency, a programmable delay is provided by tuning the back-gate (bulk) voltage of the pMOS devices of the inverters in the ring. In this way no additional loading is applied, thus maintaining the highest possible oscillator frequency.

III. DESIGN OF A UWB SC N-PATH BAND-PASS FILTER

In order to ensure proper operation, post-layout simulation of the tunable SC N-path band-pass filter depicted in Fig. 5 is provided. Simulated results of the filter for TSMC 90 nm CMOS are achieved using the Cadence design environment. The multiphase clock generator generates a 6-phase clock used for driving SC circuits. The bandwidth of the band-pass filter is shaped by the duty cycle of the clock and the low-pass filter cut-off frequency. Fig. 6 shows an example of a 6-phase clock generated by the proposed multiphase clock generator including extracted parasitic components. In this technology, operation center frequency up to 4.5 GHz can be expected.

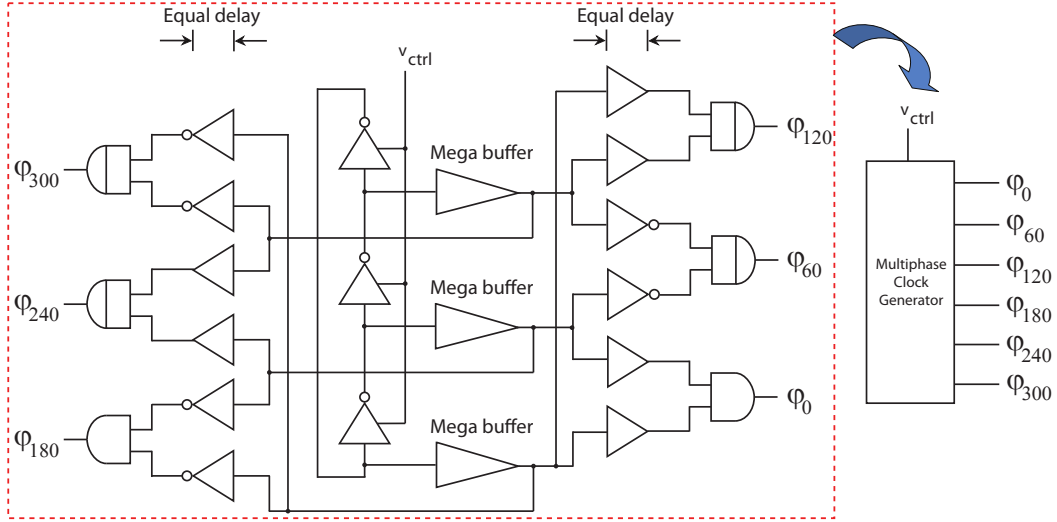


Fig. 4. The proposed multiphase clock generator.

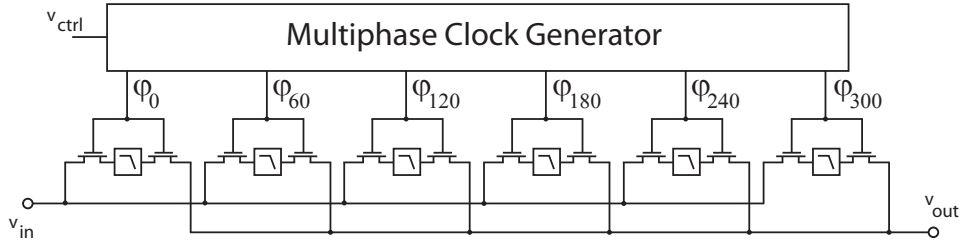


Fig. 5. The proposed UWB SC N-path band-pass filter.

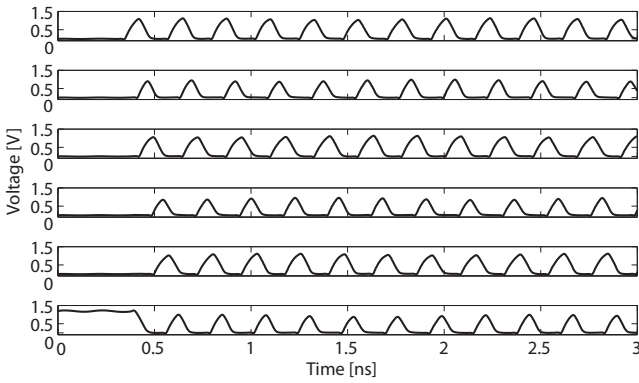


Fig. 6. An example of a 6-phase clock generated by the proposed multiphase clock generator.

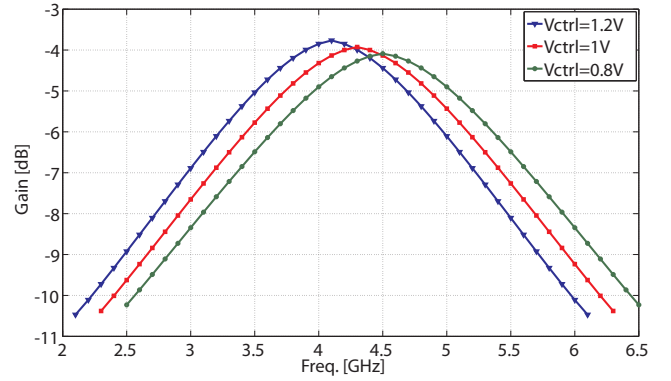


Fig. 7. The filter response at different center frequencies.

Tuning the back-gate bias voltage of the pMOS devices in the ring oscillator results in a frequency tuning range of 9.3% around 4.3 GHz. The filter responses at different center frequencies are shown in Fig. 7. Post-layout simulation shows a noise figure from 7 dB to 8 dB as presented in Fig. 8.

The wideband filter consumes a power of 35 mW from a 1.2 V supply voltage. The layout is shown in Fig. 9. It occupies an area of $0.085 \times 0.085 \text{ mm}^2$ including on-chip decoupling

capacitors. Table I summarizes the performance of the filter and compares it to other published designs in silicon. This filter offers the smallest area and highest center frequency of all the filters reported in table I.

IV. CONCLUSION

In this paper, we have presented an inductorless SC 6-path band-pass filter targeted for low band of UWB spectrum from 3.1 GHz to 5.6 GHz. As a part of the filter a novel multiphase

TABLE I
COMPARISON WITH THE STATE OF THE ART

Parameter	[1]	[3]	[9]	[10]	[11]	[12]	This work
CMOS technology	0.35 μm	65 nm	0.18 μm	0.35 μm	0.5 μm	0.5 μm	90 nm
Die area	1.9 mm^2	0.07 mm^2	0.81 mm^2	0.1 mm^2	0.15 mm^2	2.5 mm^2	0.007 mm^2
Power consumption	63 mW	2 – 16 mW	17 mW	5 mW	43.2 mW	15 mW	35 mW
Voltage Gain	-2 dB	-2 dB	0 dB	-5 dB	-15 dB	23 dB	-4 dB
-3 dB bandwidth	1.75 – 4.6 MHz	35 MHz	130 MHz	53.8 MHz	80 MHz	70 MHz	2.5 GHz
Freq. tuning range	240 – 530 MHz	0.1 – 1 GHz	2 – 2.06 GHz	1.93 – 2.19 GHz	1.77 – 1.86 GHz	2.45 – 2.85 GHz	4.1 – 4.5 GHz
Noise figure	9 dB	3 – 5 dB	15 dB	26.8 dB	28 dB	6 dB	7 – 8 dB

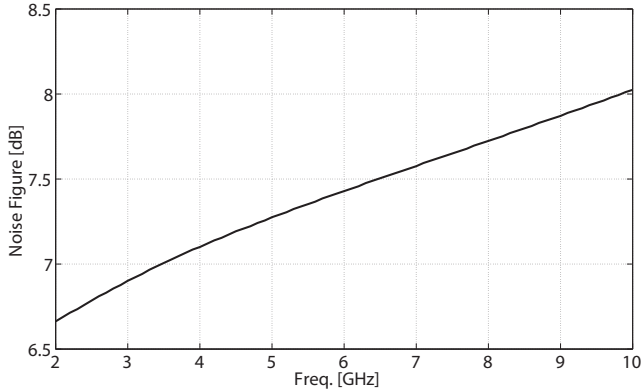
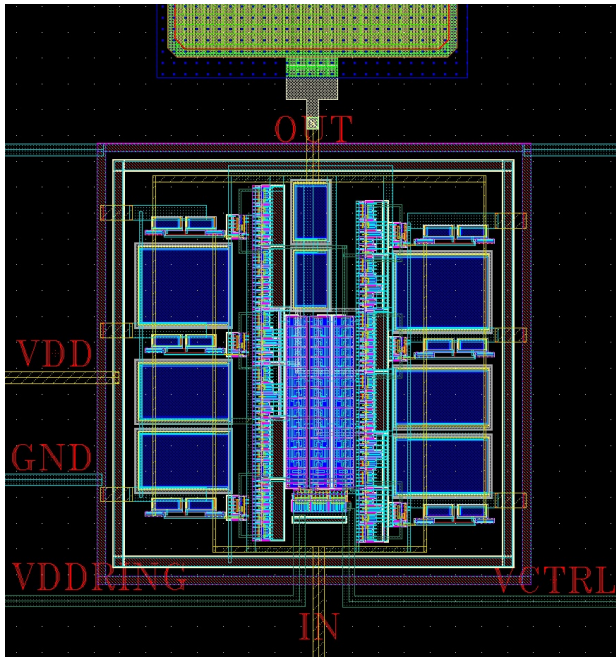


Fig. 8. The noise figure of the band-pass filter.

Fig. 9. The filter layout: $0.085 \times 0.085 \text{ mm}^2$.

clock generator is proposed suitable for RF frequencies. The band-pass filter achieves a -3 dB bandwidth of 2.5 GHz with center frequency tuning range between 4.1 GHz and 4.5 GHz. The filter is compact, suitable for standard digital technology, and can be fully integrated without any external clock reference.

ACKNOWLEDGMENT

The research has been sponsored by Norwegian Research Council project number 187857/S10 and was carried out at the Nanoelectronics group, Department of Informatics, University of Oslo, Norway.

REFERENCES

- [1] A. E. Oualkadi, M. E. Kaamouchi, J.-M. Paillot, D. Vanhoenacker-Janvier, and D. Flandre, "Fully integrated high-q switched capacitor bandpass filter with center frequency and bandwidth tuning," *2007 IEEE Radio Frequency Integrated Circuits Symposium*, pp. 681–684, 2007.
- [2] A. Mirzaei and H. Darabi, "Analysis of imperfections on performance of 4-phase passive-mixer-based high-q bandpass filters in saw-less receivers," *IEEE Transactions on Circuit and Systems-I*, vol. 58, no. 5, pp. 879–892, May 2011.
- [3] A. Ghaffari, E. A. M. Klumperink, M. C. M. Soer, and B. Nauta, "Tunable high-q n-path band-pass filters: Modeling and verification," *IEEE Journal of Solid-State Circuits*, vol. 46, no. 5, pp. 998–1010, May 2011.
- [4] L. E. Franks and F. J. Witt, "Solid-state sampled data band-pass filters," *1960 International Solid-State Circuits Conference*, pp. 70–71, February 1960.
- [5] A. W. Buchwald, K. W. Martin, A. K. Oki, and K. W. Kobayashi, "A 6-ghz integrated phase-locked loop using algaas/gaas heterojunction bioplar transistor," *IEEE Journal of Solid-State Circuits*, vol. 27, no. 2, pp. 1752–1762, December 1992.
- [6] D. C. V. Grunigen, R. P. Sigg, J. Schmid, G. S. Moschytz, and H. Melchior, "An integrated cmos switched-capacitor bandpass filter based on n-path and frequency-sampling principles," *IEEE Journal of Solid-State Circuits*, vol. 18, no. 6, pp. 753–761, December 1983.
- [7] B. Razavi, *Design of Analog CMOS Integrated Circuits*, ser. Electrical and Computer Engineering. McGraw-Hill, 2001.
- [8] T. H. Lee, *The Design of CMOS Radio-Frequency Integrated Circuits, Second Edition*. Cambridge University Press, 2004.
- [9] B. Georgescu, I. G. Finvers, and F. Ghannouchi, "A 2 ghz q-enhanced active filter with low passband distortion and high dynamic range," *IEEE Journal of Solid-State Circuits*, vol. 41, no. 9, pp. 2029–2039, September 2006.
- [10] F. Dülger, E. Sánchez-Sinencio, and J. Silva-Martínez, "A 1.3-v 5-mw fully integrated tunable bandpass filter at 2.1 ghz in 0.35-um cmos," *IEEE Journal of Solid-State Circuits*, vol. 38, no. 6, pp. 918–928, June 2003.
- [11] A. N. Mohieldin, E. Sánchez-Sinencio, and J. Silva-Martínez, "A 2.7-v 1.8-ghz fourth-order tunable lc bandpass filter based on emulation of magnetically coupled resonators," *IEEE Journal of Solid-State Circuits*, vol. 38, no. 7, pp. 1172–1181, July 2003.
- [12] X. He and W. B. Kuhn, "A 2.5-ghz low-power, high dynamic range, self-tuned q-enhanced lc filter in soi," *IEEE Journal of Solid-State Circuits*, vol. 40, no. 8, pp. 1618–1628, August 2005.

Paper 24 Continuous-time CMOS quantizer for ultra-wideband applications

T. A. Vu, S. Sudalaiyandi, M. Z. Dooghabadi, H. A. Hjortland, O. Nass, T. S. Lande, and S. E. Hamran. “Continuous-time CMOS quantizer for ultra-wideband applications”. In: *Circuits and Systems (ISCAS), Proceedings of 2010 IEEE International Symposium on*, pp. 3757–3760, Paris, France, May 2010. doi:10.1109/ISCAS.2010.5537745. © 2010 IEEE. Reprinted with permission.

My Contributions to This Paper

The challenging element in a CTBV processing system is the single-bit quantizer. For swept-threshold functionality we need a comparator with high gain. However, most high-speed comparators are clocked and cannot be used. This paper is an effort to find a wideband and continuous-time comparator suitable for CTBV applications. I contributed to the development.

Continuous-Time CMOS Quantizer for Ultra-Wideband Applications

Tuan Anh Vu, Shanthi Sudalaiyandi, Malihe Zarre Dooghabadi,
Håkon A. Hjortland, Øivind Næss, Tor Sverre Lande and Svein Erik Hamran
Dept. of Informatics, University of Oslo, Norway
Email: anhtv@ifi.uio.no

Abstract—In this paper, a novel continuous-time CMOS quantizer suitable for ultra-wideband (UWB) single-bit quantization is presented. The proposed quantizer achieves a -3 dB bandwidth covering the entire FCC UWB spectrum from 3.1 GHz to 10.6 GHz with a very high gain of approximately 70 dB. The quantizer is an area-efficient, single-inductor solution designed for TSMC 90 nm CMOS technology with the supply voltage of 1.2 V.

I. INTRODUCTION

As ultra wideband, wireless applications are emerging [1], implementations in standard CMOS technology is an important factor for cost-efficient production. Although nanometer technology provide fast devices, it is hard to find solutions in pure CMOS for the FCC released UWB band of 3.1 GHz to 10.6 GHz. Apparently clock frequencies in PCs are saturating around 3 GHz.

At the Nanoelectronics group, Dept. of Informatics, Univ. of Oslo we have proposed a design paradigm called Continuous-Time, Binary-Value (CTBV) facilitating gigahertz operation in standard CMOS technology [2] [3]. The high speed operation is obtained by eliminating clocks. A similar approach is explored for low power operation in [4] and [5]. The CTBV design paradigm has even been explored commercially for single chip impulse radar solutions [1].

A vital operation of CTBV systems is a coarse binary quantizer or thresholder continuously comparing the incoming signal to a threshold voltage, giving a binary output of the sign of the comparison. Such a function may be obtained with a high-gain comparator. Most high-speed comparators are using a clocked reset [6], but such an approach cannot be used for CTBV solutions. An important consideration is the DC gain requirement after the thresholding operation, limiting viable circuit solutions. Another observation is the high gain requirement for correct quantization of small signal amplitudes from weak transmissions.

In this paper we are proposing a solution for continuous-time, high-gain quantizer suitable for ultra wideband applications. A bandwidth exceeding 10 GHz is feasible while maintaining sufficient DC gain for the thresholding operation. The proposed solution is designed in 90nm TSMC technology exploring resistive-feedback inverters and a single LC resonator at the input.

II. QUANTIZER CIRCUIT

The proposed quantizer circuit is shown in Fig. 1. It consists of cascaded amplifier stages with an input LC resonator for extended bandwidth. The threshold voltage V_{th} is mixed with the input signal at the amplifier input providing a controllable quantization voltage level for input signal thresholding. For improved noise performance the threshold voltage is set by a current I_{th} and converted to a suitable threshold voltage level matching the build-in inverter threshold.

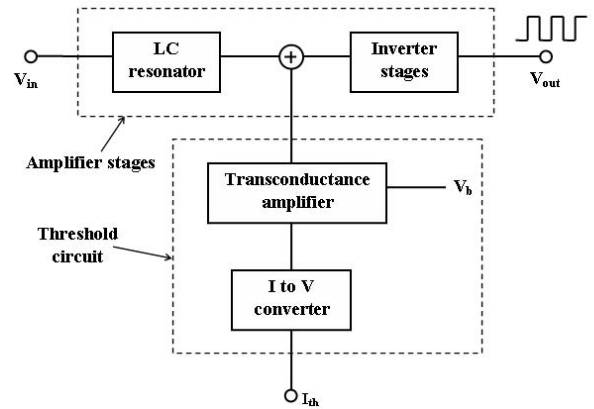


Fig. 1. The structure of the proposed quantizer.

A. Amplifier Stages

A simple digital inverter with resistive feedback is used for each amplifier stage. For increased bandwidth, strong feedback is applied sacrificing stage gain. In order to recover gain, a cascade of inverting amplifiers are used as shown in Fig. 2. Wider bandwidth is achieved at the expense of lower gain per stage by using low values of R_f . Amplifier bandwidth is mainly determined by RC time constants of every node. In CMOS technology, severe parasitic capacitance deteriorates bandwidth significantly. Considering the inter-stage small-signal model, the transfer function can be expressed as [7]:

$$\frac{V_{out}}{V_{in}} = \frac{-g_m R_T}{1 + s C_T R_T}$$

Where R_T denotes $R_{f1} || R_{f2}$ and C_T represent $C_1 + C_2$. R_{f1}/R_{f2} and C_1/C_2 denote equivalent resistors and capacitors contributed by previous and next stages, respectively.

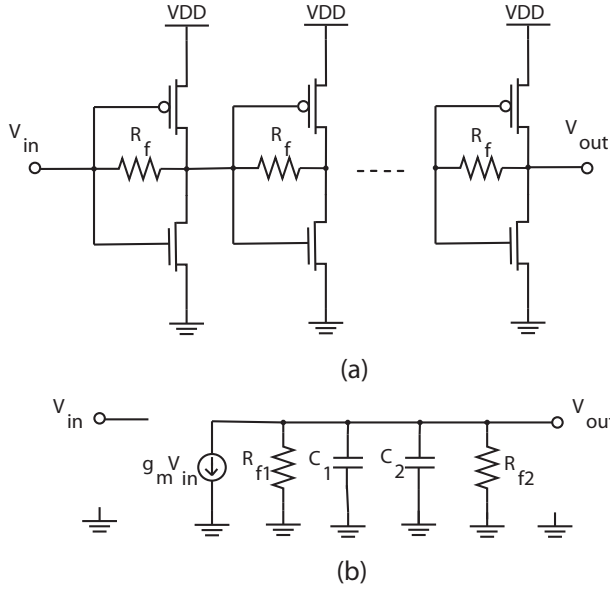


Fig. 2. (a) Resistive-feedback inverter stages. (b) Equivalent inter-stage small-signal model.

However, using resistive feedback is the simplest type of broadbanding. It is also the most inefficient: Lower gain, lower output power, and degraded noise figure [8]. In order to obtain high bandwidth, conventional multiple inductive-series peaking technique is used in [7] which is demonstrated in Fig. 3a. In this architecture inductors are deployed between each amplifier. Together with parasitic capacitances a third-order LC-ladder filter is giving an impedance transformation network [9] [10]. Another structure with some improvement is introduced in [11] and depicted in Fig. 3b. By locating a peaking inductor at the gate of nMOS of each inverter stage, the -3 dB roll-off frequency can be boosted to higher frequencies. However, using peaking inductor for bandwidth extension in each gain stage is area demanding.

In this paper, a novel structure for the wideband amplifier is proposed. Instead of using multiple peaking inductors, a resonant peak at the amplifier corner frequency can 'pull up' the gain, thus extending the bandwidth significantly as shown in Fig. 5. In this design, only a single, small inductor (0.82 nH) is used for the LC resonator regardless of the number of amplifier stages. The LC resonator also acts as a high-pass filter at the input, shifting the bandwidth to higher frequencies suitable for the FCC approved UWB spectrum. Table I gives a comparison of some reported designs. The proposed amplifier architecture is shown in Fig. 4.

B. Threshold Circuit

Signal and threshold level are added together at the input of the quantizer. In order to reduce noise, a threshold current, I_{th} , is used as input. I_{th} is then converted into a suitable threshold voltage, V_{th} by a simple I-V converter followed by a transconductance amplifier in a voltage follower configuration as shown in Fig. 6. The I-V converter acts as a linear circuit

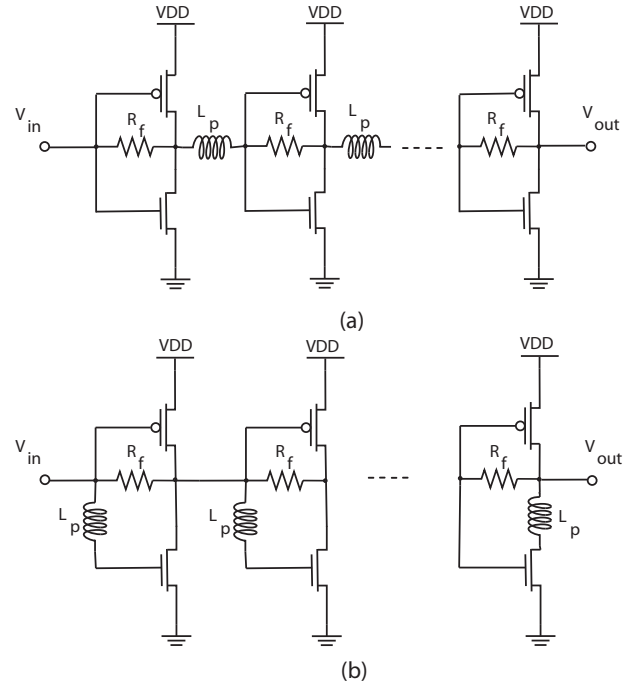


Fig. 3. (a) Conventional multiple inductive-series peaking technique. (b) Splitting-load inductive peaking technique.

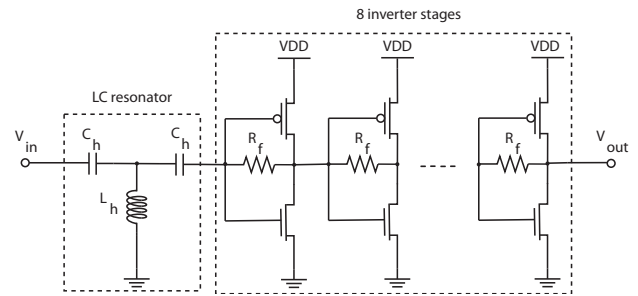


Fig. 4. The proposed high-gain UWB amplifier.

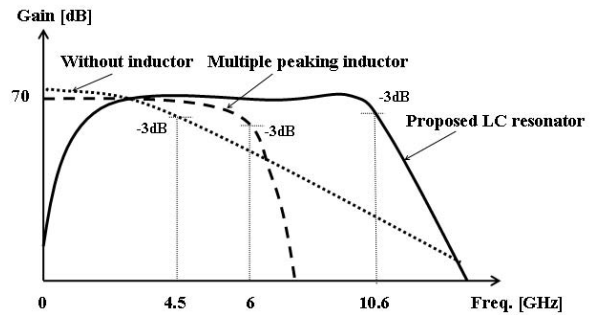


Fig. 5. Bandwidth comparison among the designs.

TABLE I
COMPARISON WITH THE STATE OF THE ART

Design	TIA [7]	MMIC [11]	This Work
CMOS technology	0.18 μm	0.13 μm	90 nm
Supply voltage	1.8 V	1.3 V	1.2 V
Gain (dB)	61	13.2	70
-3 dB bandwidth	DC–7.2 GHz	DC–1.5 GHz	3.1 GHz–10.6 GHz
No. of stages	5	3	8
No. of inductors	8 (1.1 nH)	3 (2.4 nH, 2.4 nH, and 1.4 nH)	1 (0.82 nH)

with transfer ratio $k = V_{th}/I_{th}$. By changing V_b the tail current of the transconductance amplifier, offset strength may be changed in the mixing node as well as changing the corner frequency of the high-pass filter action following the input resonator.

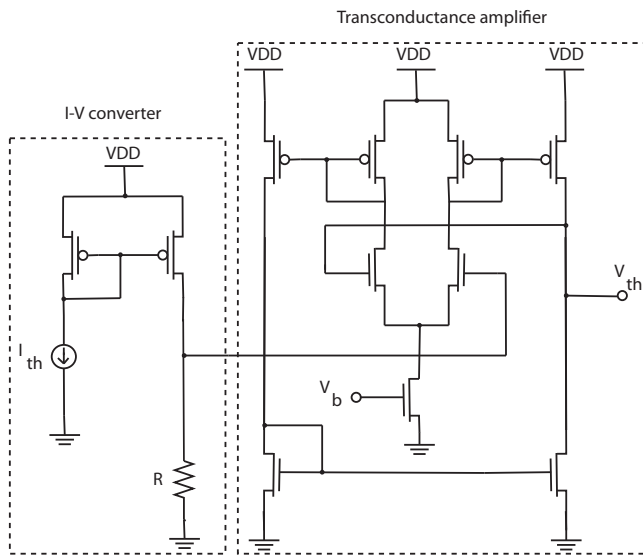


Fig. 6. The threshold circuit.

III. SIMULATED RESULTS

Simulated results of the quantizer for TSMC 90 nm CMOS technology is achieved using the CADENCE design environment. Circuit design at high frequencies involves more detailed considerations than at lower frequencies when the effect of parasitic capacitances and inductances can impose serious constraints on achievable performance. Thus, all components used for simulation are RF models provided by TSMC. The complete circuit for the proposed quantizer with all component values are given in Fig. 9.

In Fig. 7 the expected performance of the threshold tuning circuit is shown. The threshold voltage change linearly with the applied input current when varying from zero to 100 μA . Then it becomes saturated as the current continues increasing. The range for threshold voltage, V_{th} , is between GND and 800 mV. Simulated results also shows that the operating point of the quantizer is around 520 mV so the threshold level is from -520 mV to 380 mV.

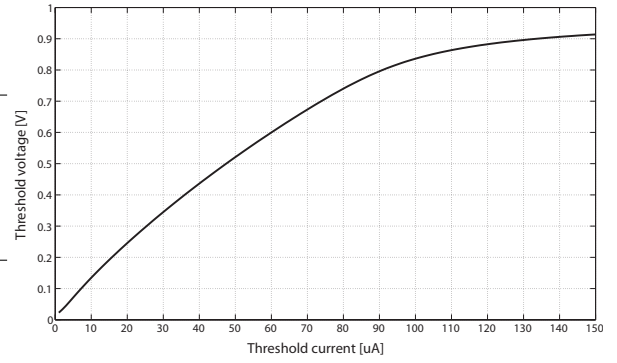


Fig. 7. The performance of the threshold circuit.

Fig. 8 shows the frequency response of the quantizer. A high gain of approximately 70 dB throughout the entire UWB spectrum, 3.1 GHz–10.6 GHz, is demonstrated.

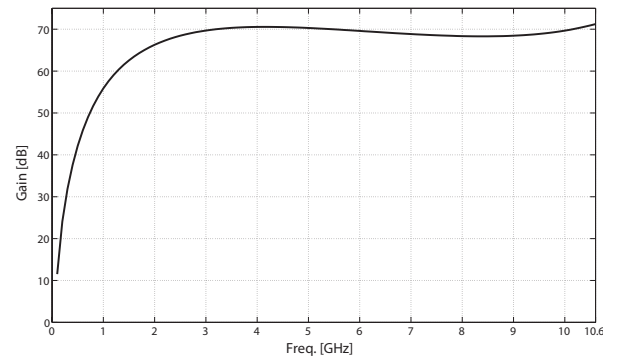


Fig. 8. The frequency response of the quantizer.

Fig. 10 gives the performance of the quantizer when a sinusoidal signal is applied at the input for some threshold levels. The input signal has the amplitude of 2.5 mV, frequency of 5 GHz. The threshold voltage is swept with the step size of 100 μV .

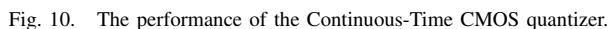
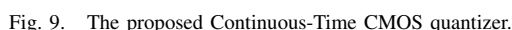
As required, even minor signal perturbations are quantized.

IV. CONCLUSION

In this paper we have proposed a continuous-time, ultra wideband quantizer with tunable threshold level and high gain suitable for FCC UWB applications. The quantizer is intended for CTBV architectures with low-power operation. Area-efficiency is good using only one inductor. Compared to other reported results, we have superior performance.

ACKNOWLEDGMENT

The research has been sponsored by Norwegian Research Council project number 187857/S10 and was carried out at the Nanoelectronics group, Department of Informatics, University of Oslo, Norway.



- ## REFERENCES
- [1] [Online]. Available: <http://www.novelda.no/>
 - [2] H. A. Hjortland and T. S. Lande, "CTBV integrated impulse radio design for biomedical applications," *IEEE Transactions on Biomedical Circuits and Systems*, vol. 3, no. 2, pp. 79–88, Apr. 2009.
 - [3] H. A. Hjortland, D. T. Wisland, T. S. Lande, C. Limbodal, and K. Meisal, "Thresholded samplers for uwb impulse radar," in *IEEE International Symposium on Circuits and Systems, 2007. ISCAS 2007.*, 2007.
 - [4] Y. Li, K. Shepard, and Y. Tsividis, "A continuous-time programmable digital fir filters," *IEEE Journal of Solid-State Circuits*, vol. 41, no. 11, pp. 2512–2530, Nov. 2006.
 - [5] B. Schnell and Y. Tsividis, "A continuous-time adc/dsp/dac system with no clock and with activity-dependent power dissipation," *IEEE Journal of Solid-State Circuits*, vol. 43, no. 11, pp. 2472–2481, Nov. 2008.
 - [6] B. Goll and H. Zimmermann, "A 65nm cmos comparator with modified latch to achieve 7ghz/1.3mw at 1.2v and 700mhz/47uw at 0.6v," pp. 328–329, February 2009.
 - [7] C.-H. Wu, C.-H. Lee, W.-S. Chen, and S.-I. Liu, "Cmos wideband amplifiers using multiple inductive-series peaking technique," *IEEE Journal of Solid-State Circuit*, vol. 40, no. 2, pp. 548–552, February 2005.
 - [8] R. Goyal, "High-frequency analog integrated circuit design," in *Willey Series in Microwave and Optical Engineering*, 1995.
 - [9] R. Schaumann and M. Valkenburg, "Design of analog filters," in *New York: Oxford Univ. Press*, 2001.
 - [10] S. Galal and B. Razavi, "A 40gb/s amplifier end esd protection circuit in 0.18um cmos technology," in *IEEE International Solid-State Circuits Conference (ISSCC) Dig. Tech. Papers*, February 2004, pp. 480–481.
 - [11] S.-F. Chao, J.-J. Kuo, C.-L. Lin, M.-D. Tsai, and H. Wang, "A dc-11.5ghz low-power, wideband amplifier using splitting-load inductive peaking technique," *IEEE Microwave and wireless components letters*, vol. 18, no. 7, pp. 482–484, July 2008.

Paper 25 A 70-dB, 3.1-10.6-GHz CMOS amplifier in low-power 90 nm CMOS

T. A. Vu, S. Sudalaiyandi, H. A. Hjortland, O. Naess, T. S. Lande, and S. E. Hamran. “A 70-dB, 3.1-10.6-GHz CMOS amplifier in low-power 90 nm CMOS”. In: *Circuits and Systems (MWSCAS), 2011 IEEE 54th International Midwest Symposium on*, pp. 1–4, Seoul, Aug. 2011. doi:10.1109/MWSCAS.2011.6026639. © 2011 IEEE. Reprinted with permission.

My Contributions to This Paper

This paper presents a high-frequency high-gain amplifier. I took part in discussions about the source-follower buffer and I came up with the design for the test circuits which can be used to characterize the amplifier and which are shown in the layout figure.

A 70-dB, 3.1-10.6-GHz CMOS Amplifier in Low-Power 90 nm CMOS

Tuan Anh Vu, Shanthi Sudalaiyandi, Håkon A. Hjørtland,
Øivind Næss, Tor Sverre Lande and Svein Erik Hamran
Dept. of Informatics, University of Oslo, Norway
Email: anhtv@ifi.uio.no

Abstract—This paper presents a novel high-gain CMOS amplifier suitable for ultra-wideband (UWB) applications. The proposed amplifier achieves a -3 dB bandwidth covering the entire FCC UWB spectrum from 3.1 GHz to 10.6 GHz with a very high gain of approximately 70 dB. The amplifier is an area-efficient, single-inductor solution designed for TSMC 90 nm CMOS low-power process while consuming 25.1 mW from 1.2 V supply voltage.

I. INTRODUCTION

Ultra-wideband transmission has recently received significant attention in both academic and industry for applications in wireless communications [1][2]. UWB has many benefits, including high data rate, availability of low-cost transceiver, low transmit power and low interference. The exploration of UWB technology has forced innovative circuit design in order to achieve sufficient performance in standard digital CMOS technology, especially when low-power solutions are desired. CMOS amplifiers are the main building blocks in many analog signal processing systems as well as other analog interfacing circuits. Also in communication equipment, signal acquisition amplification is required. With microwave frequencies in combination with wide bandwidth, amplifier design is getting even more challenging. The design of UWB amplifiers in CMOS technology requires careful design. Major challenges are the limited power supply and significant parasitics.

In this paper, the analysis, design and implementation of a novel high-gain UWB amplifier is proposed. The amplifier is based on cascaded multistage resistive-feedback inverters with LC resonator for bandwidth extension. This is confirmed by post-layout simulation in a TSMC 90nm CMOS low-power process. Simulated results show that this amplifier exhibits a 70 dB gain throughout the bandwidth from 3.1 GHz to 10.6 GHz.

II. CIRCUIT DESCRIPTION

The proposed amplifier is using standard “digital” inverters combined with passive devices and cascaded for sufficient overall gain.

A. Amplifying Stages

A simple digital inverter with resistive feedback is used for each amplifying stage. A very high gain can be achieved by cascading multiple inverter stages. However, there is a trade-off between the gain and the bandwidth of the amplifier.

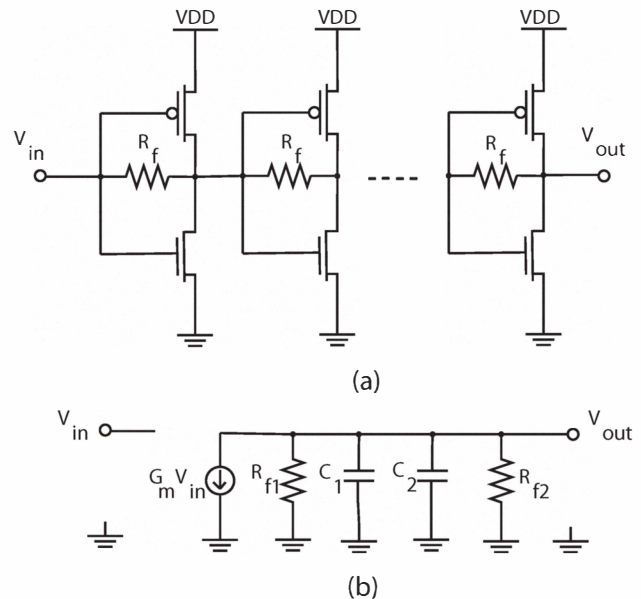


Fig. 1. (a) Resistive-feedback inverter stages. (b) Equivalent inter-stage small-signal model.

For increased bandwidth, stage gain is sacrificed for strong feedback. In order to recover gain, a cascade of inverting amplifiers is used as shown in Fig. 1. Wider bandwidth is obtained at the expense of lower gain per stage by using low values of R_f . Amplifier bandwidth is mainly determined by RC time constants of each node. In CMOS technology, parasitic capacitance reduces bandwidth significantly. Considering the inter-stage small-signal model, the transfer function can be expressed as [3]:

$$\frac{V_{out}}{V_{in}} = \frac{-G_m R_T}{1 + s C_T R_T}$$

Where R_T denotes $R_{f1} || R_{f2}$ and C_T represent $C_1 + C_2$. R_{f1}/R_{f2} and C_1/C_2 denote equivalent resistors and capacitors contributed by previous and next stages, respectively. However, resistive feedback is the simplest type of broadbanding. It is also the most inefficient: reduced gain, reduced output power, and degraded noise figure [4]. In order to obtain high bandwidth, conventional multiple inductive-series peaking technique is used as proposed in [3] and is demonstrated in Fig. 2a. In this architecture inductors are deployed between

each amplifier. Together with parasitic capacitances a third-order LC-ladder filter is giving an impedance transformation network [5][6]. Another structure with some improvement is introduced in [7] and depicted in Fig. 2b. By locating a peaking inductor at the gate of nMOS of each inverter stage, the -3 dB roll-off frequency can be boosted to higher frequencies. However, using peaking inductor for bandwidth extension in each gain stage is area demanding.

A novel structure for the wideband amplifier is proposed. Instead of using multiple peaking inductors, a resonant peak at the amplifier corner frequency can 'pull up' the gain, thus extending the bandwidth significantly. In this design, only a single, small inductor (0.7 nH) is used for the LC resonator regardless of the number of amplifier stages. The LC resonator also acts as a high-pass filter at the input, shifting the bandwidth to higher frequencies suitable for the FCC approved UWB spectrum. The proposed amplifier architecture is shown in Fig. 4. It is the core of the continuous-time CMOS quantizer introduced in [8] but the amplifier presented here may find usage in other wideband applications. A very high gain is required for high input sensitivity quantization thus, a very small signal can be detected.

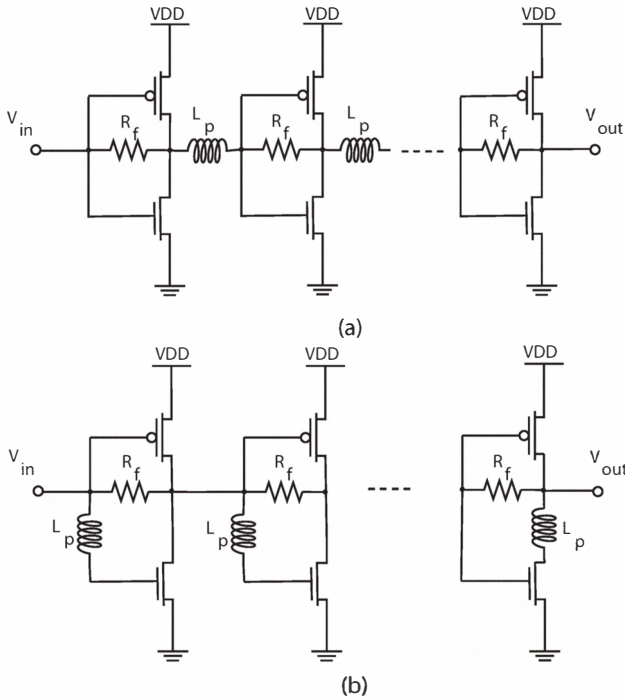


Fig. 2. (a) Conventional multiple inductive-series peaking technique. (b) Splitting-load inductive peaking technique.

B. Impedance Matching Circuit

Since the LC resonator produces very low impedance at the input, a buffer is required to drive the source with negligible signal loss. A source follower configuration is proposed for this purpose. A current mirror supplies the bias current for the source-follower. Although the DC level of the output voltage

TABLE I
TRANSISTOR DIMENSIONS, COMPONENT VALUES AND BIAS SETTING

Parameter	Value
VDD	1.2 V
I_{bias}	10 mA
V_{bias}	500 mV
M_1, M_2, M_3	80 $\mu\text{m}/110\text{ nm}$
M_4, M_6, \dots, M_{26}	5 $\mu\text{m}/100\text{ nm}$
M_5, M_7, \dots, M_{27}	4.08 $\mu\text{m}/100\text{ nm}$
C_{h1}	0.5 pF
C_{h2}	0.7 pF
L_h	0.7 nH
R	50 Ω
R_f	7.8 k Ω

is somewhat below the DC level of the input voltage, ideally the small-signal voltage gain is close to unity. In reality, it is somewhat less than unity [9]. By changing V_{bias} and I_{bias} the high input impedance and moderate output impedance of the buffer can be tunable. The output impedance is mainly decided by the cascaded inverter stages requiring no additional buffering.

III. SIMULATED RESULTS

Simulated results of the amplifier for TSMC 90 nm CMOS low-power process is achieved using the CADENCE design environment. Circuit design at high frequencies involves more detailed considerations than at lower frequencies when the effect of parasitic capacitances and inductances can impose serious constraints on achievable performance. Thus, all components used for simulation are RF models provided by TSMC. Transistor dimensions, component values and bias setting are shown in Table I.

The performance of the source-follower buffer is demonstrated in Fig. 3. As shown in the figure, the attenuation is negligible while maintaining a flat frequency response throughout the desired frequency range between 3.1 GHz and 10.6 GHz. This ensures the efficient use of the source-follower buffer at the input.

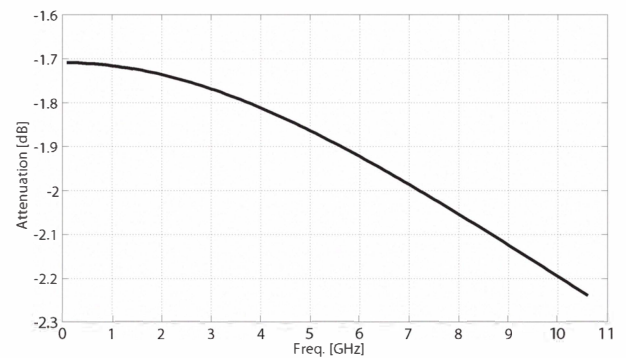


Fig. 3. The frequency response of the source-follower buffer.

The input impedance of the amplifier is designed for matching to 50 Ω . Fig. 5 and Fig. 6 shows the simulated S_{11} in dB versus frequency and a corresponding amplifier's input

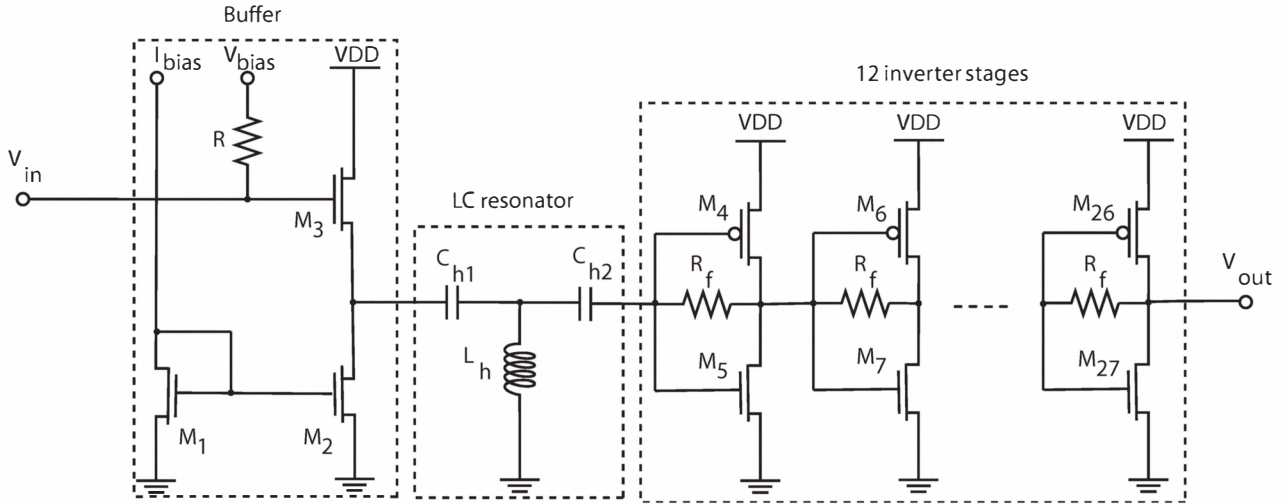


Fig. 4. The proposed high-gain UWB amplifier.

impedance seen in the Smith chart. As depicted in Fig. 5 and Fig. 7 the amplifier presents good input matching better than -18 dB over the entire UWB spectrum. In most of frequency band, the voltage standing wave ratio (VSWR) is less than 1.6 indicating that the reflection coefficient is less than 0.2 across the desired frequency range. Fig. 6 shows the input impedance of the amplifier from 3.1 GHz to 10.6 GHz. As can be seen in the Smith chart, the input impedance is close to the center where it is perfectly matched to 50 Ω .

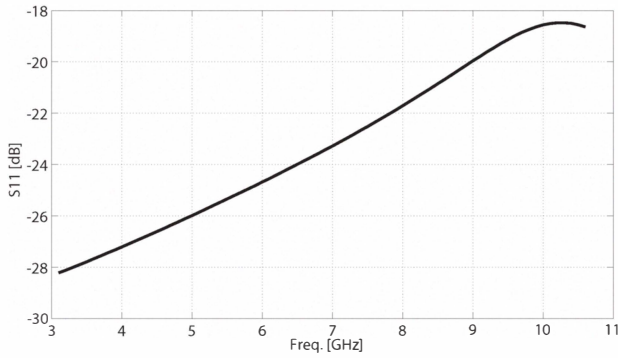


Fig. 5. The input return loss of the amplifier.

The Rollett's stability factor is calculated over the frequency band 3.1 GHz–10.6 GHz by using the below equation [10]:

$$K = \frac{1 + |S_{11}S_{22} - S_{12}S_{21}|^2 - |S_{11}|^2 - |S_{22}|^2}{2|S_{12}S_{21}|}$$

Since its value is always greater than unity, the circuit claims unconditional stability. Besides, the noise performance can be improved significantly by feeding the high-gain amplifier with a signal from a low noise amplifier (LNA) as often found in wireless front-ends.

Fig. 8 gives the value of the gain obtained for the entire band of interest. The gain is maintained approximately

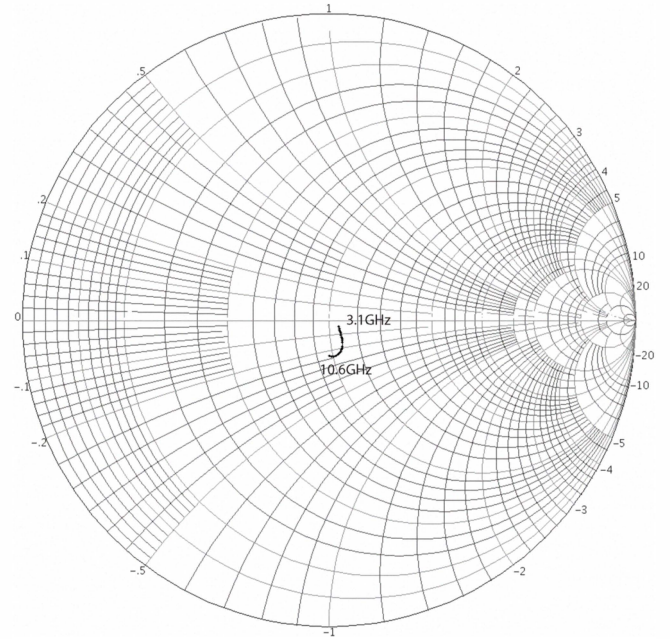


Fig. 6. The input impedance from 3.1 GHz to 10.6 GHz seen in the Smith chart.

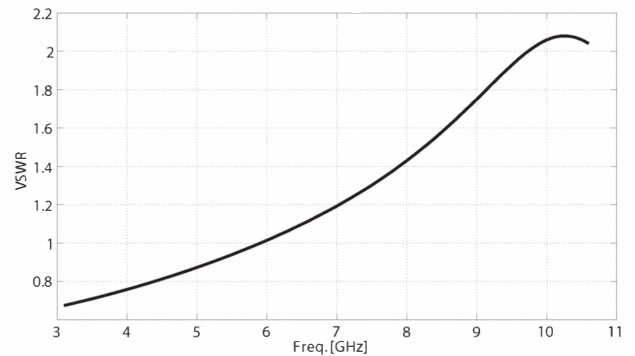


Fig. 7. The input VSWR of the amplifier.

TABLE II
COMPARISON WITH THE STATE OF THE ART

Design	TIA [3]	MMIC [7]	This Work
CMOS tech.	0.18 μm	0.13 μm	90 nm
Power cons.	70 mW	9.1 mW	25.1 mW
Gain (dB)	61	13.2	70
-3 dB BW	DC–7.2 GHz	DC–11.5 GHz	3.1 GHz–10.6 GHz
No. of stages	5	3	12
No. of inductors	8 (1.1 nH)	3 (2.4 nH, 2.4 nH and 1.4 nH)	1 (0.7 nH)

70 dB, including approximately -2 dB attenuation from the source-follower stage, throughout the entire UWB spectrum, 3.1 GHz–10.6 GHz. The peak gain of 70.7 dB is achieved at 4 GHz. Fig. 9 shows the linearity of the amplifier in terms of 1 dB compression point when a 4 GHz-sinusoidal signal is applied at the input. The layout of the amplifier is shown in Fig. 10. The proposed amplifier occupies a die area of 0.32 x 0.38 mm². Table II gives a comparison of some reported designs.

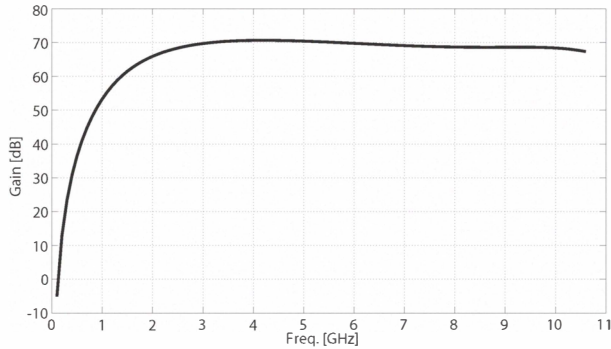


Fig. 8. The frequency response of the amplifier.

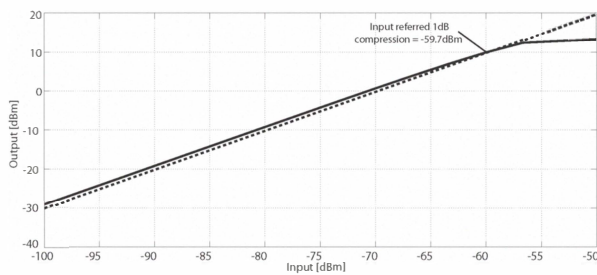


Fig. 9. The linearity of the amplifier in terms of 1 dB compression point.

IV. CONCLUSION

This paper proposes a cascaded high-gain amplifier exploring inverters with resistive feedback. The proposed amplifier is suitable for ultra-wideband microwave signals and is area-efficient with only one inductor. An implementation in TSMC 90 nm CMOS technology indicates approximately 70 dB gain. Compared to similar published topologies, our proposed amplifier is expected to be superior.

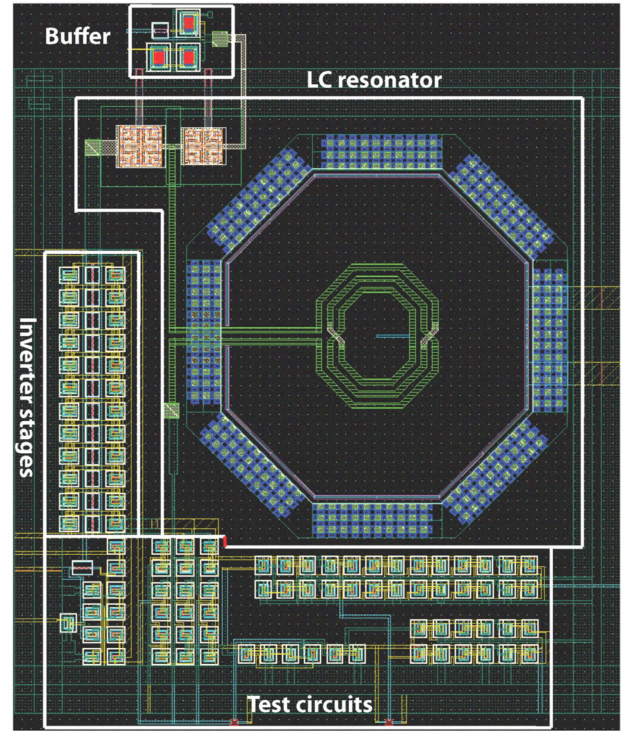


Fig. 10. The layout of the proposed amplifier.

ACKNOWLEDGMENT

The research has been sponsored by Norwegian Research Council project number 187857/S10 and was carried out at the Nanoelectronics group, Department of Informatics, University of Oslo, Norway.

REFERENCES

- [1] J. R. Foerster, "The performance of a direct-sequence spread ultra wide-band system in the presence of multipath, narrowband interference, and multiuser interference," *Proc. Conference on Ultra-Wideband Systems and Technologies*, pp. 87–92, May 2002.
- [2] B. M. Sadler and A. Swami, "On the performance of uwband and ds-spread spectrum communication systems," *Proc. Conference on Ultra-Wideband Systems and Technologies*, pp. 289–292, May 2002.
- [3] C.-H. Wu, C.-H. Lee, W.-S. Chen, and S.-I. Liu, "Cmos wideband amplifiers using multiple inductive-series peaking technique," *IEEE Journal of Solid-State Circuit*, vol. 40, no. 2, pp. 548–552, February 2005.
- [4] R. Goyal, "High-frequency analog integrated circuit design," *Wiley Series in Microwave and Optical Engineering*, 1995.
- [5] R. Schaumann and M. Valkenburg, "Design of analog filters," *New York: Oxford Univ. Press*, 2001.
- [6] S. Galal and B. Razavi, "A 40gb/s amplifier and esd protection circuit in 0.18um cmos technology," *IEEE International Solid-State Circuits Conference (ISSCC) Dig. Tech. Papers*, pp. 480–481, February 2004.
- [7] S.-F. Chao, J.-J. Kuo, C.-L. Lin, M.-D. Tsai, and H. Wang, "A dc-11.5ghz low-power, wideband amplifier using splitting-load inductive peaking technique," *IEEE Microwave and wireless components letters*, vol. 18, no. 7, pp. 482–484, July 2008.
- [8] T. A. Vu, S. Sudalaiyandi, M. Z. Dooghabadi, H. A. Hjortland, Øivind Næss, T. S. Lande, and S. E. Hamran, "Continuous-time cmos quantizer for ultra-wideband applications," *2010 IEEE International Symposium on Circuits and Systems (ISCAS2010)*, pp. 3757–3760, June 2010.
- [9] B. Razavi, "Design of analog cmos integrated circuits," *McGraw-Hill Series in Electrical and Computer Engineering*, 2001.
- [10] G. Gonzalez, "Microwave transistor amplifiers - analysis and design /second edition," *Prentice Hall, Inc.*, 1997.

Paper 26 Power Harvesting Circuits in 90 nm CMOS

T. K. Halvorsen, H. A. Hjortland, and T. S. Lande. “Power Harvesting Circuits in 90 nm CMOS”. In: *NORCHIP, 2008.*, pp. 154–157, Tallinn, Nov. 2008. doi : 10.1109/NORCHP.2008.4738301. © 2008 IEEE. Reprinted with permission.

My Contributions to This Paper

As co-supervisor I worked with the candidate regarding implementation and measurement setup. I also worked with the candidate on the analysis of charge pump solutions.

Power Harvesting Circuits in 90 nm CMOS

Trygve K. Halvorsen, Håkon A. Hjortland and Tor Sverre “Bassen” Lande

Dept. of Informatics, University of Oslo, Norway

E-mail: {trygvekh, haakoh, bassen}@ifi.uio.no

Abstract—The purpose of this paper is to explore power harvesting capabilities in nanometer technology. Novel charge pump improvements using the back-gate or well of MOS devices improve efficiency as well as sensitivity. The proposed circuits are implemented in 90 nm CMOS. Measured performance will be provided.

I. INTRODUCTION

Power harvesting integrated as part of a standard CMOS technology would be useful in several applications. Typically in RFID tags, inductive power harvesting is used [1] requiring fairly large, external inductors and working primarily in near field. Exploring electromagnetic waves over longer distances is tempting and might even recover radio waves from devices like mobile phone or wireless LAN. To some extent the downsizing of technology is reducing power requirements, but reduced supply voltage is limiting the controllable on-chip voltages to less than 1 V. As circuits, both digital and analog, are exploring subthreshold transistor operation [2], power harvesting circuit elements for fine-pitch CMOS technology is worthwhile exploring. In this paper we will present the fundamental charge pump circuit and explore design improvements for better efficiency.

II. CHARGE PUMP CIRCUITS

The original Dickson charge pump [3] circuit is shown in Fig. 1.

The major building blocks required are capacitors and rectifying elements, i.e. diodes. As capacitors are readily available in CMOS, the challenging element is the rectifier element. The first order operation of the original Dickson charge pump is the two-phase clock that is capacitively lifting the node charge on the rising edges through the diode, while the diodes are preventing the charge to leak back on the negative clock

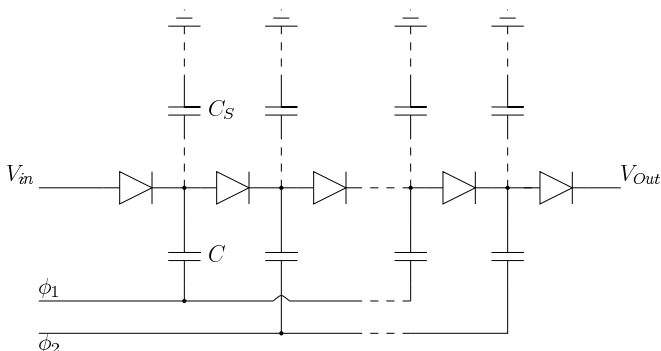


Fig. 1. The original Dickson charge pump

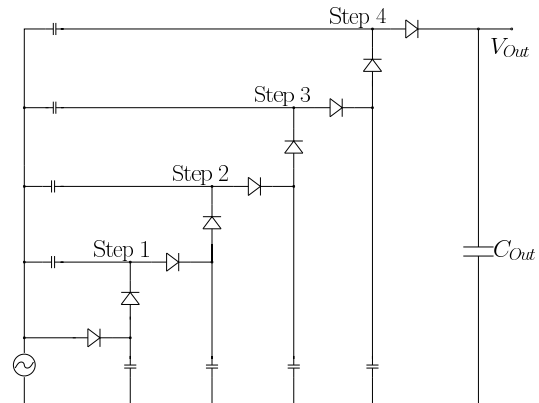


Fig. 2. Four-step cascode charge pump

transition. For power harvesting a single phase architecture is necessary like the one in Fig. 2.

A single AC source is rectified and charge is built up using capacitors and diodes [4]. Our implementation in silicon is based on the circuit topology of Fig. 2. By exploring the design space of modern technology, some reasonably efficient circuits are feasible, and in the following we will explore modern CMOS technology for improved charge pump performance.

III. CMOS TECHNOLOGY OPTIONS

The optimal rectifying element would be a device with low or no resistance when forward biased and with no conductance or infinite resistance when reverse biased. The normal diode forward biasing voltage of 0.7 V makes junction diodes unsuitable for power harvesting. Incoming antenna signals are normally <100 mV and the forward biasing voltage must be within the range of the received signal. The substitute of a junction diode in CMOS is a diode-connect MOS transistor. In addition, a number of different threshold voltages are available as different devices; low, standard and high threshold. Low gate leakage devices are also available, but often the transconductance is reduced.

As shown in Fig. 3, the differences between the threshold voltages are significant when comparing the 3 different devices of 90 nm MOS transistors. In order to make the charge pump most efficient, the threshold voltages of the diodes have to be low when pumping the voltage in order to get the voltage drop minimized. On the other hand when pushing the charge further on, at the negative half-period of the input, a high threshold is needed in order to reduce the leakage in the reverse direction of the diode as much as possible. Benefits of

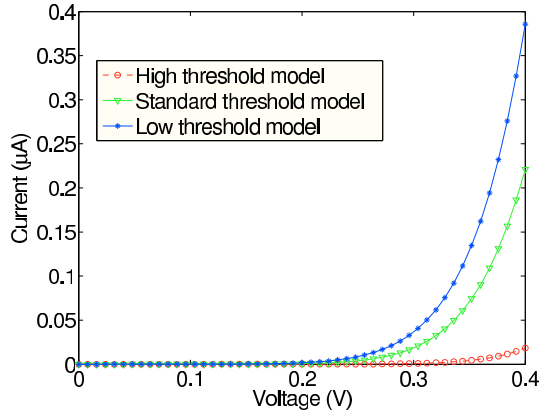


Fig. 3. MOS transistors diode-connected (low, standard and high threshold).

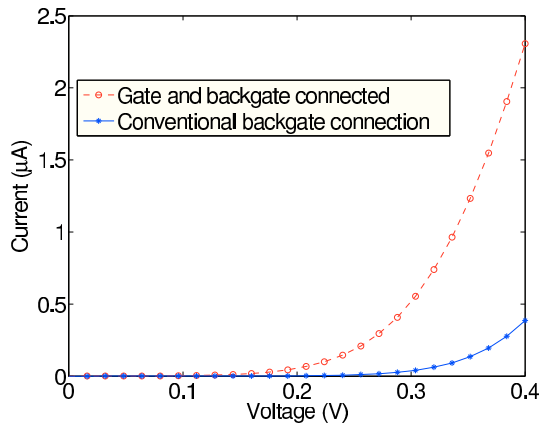


Fig. 4. A classically diode connected PMOS transistor (low threshold device) compared to one with gate and backgate connected.

increased forward current outweighs the detrimental effects of the increased reverse current.

Another option of modern CMOS processes is triple well devices. With low supply voltages, the well potential may be explored for both dynamic threshold adjustment and improved efficiency. First we can take a look at the following equation [5],

$$V_{th} = V_{th0} + \gamma \left(\sqrt{V_{sb} + |2\phi_F|} - \sqrt{|2\phi_F|} \right) \quad (1)$$

where γ is the body effect parameter, $2\phi_F$ is the surface potential, and V_{th0} is the zero bias threshold voltage. V_{sb} is a parameter that can be altered to achieve a lower V_{th} . This is done by connecting the back-gate/bulk contact to a different potential, hence causing a reversed body effect. The bulk connection can be thought of as a second gate, and can alter the threshold.

With this knowledge in mind, a charge pump based on the one in Fig. 2 was designed.

IV. CMOS HIGH SENSITIVITY CHARGE PUMP

The charge pump can be described as a series of steps, and one step of the pump is shown in Fig. 5.

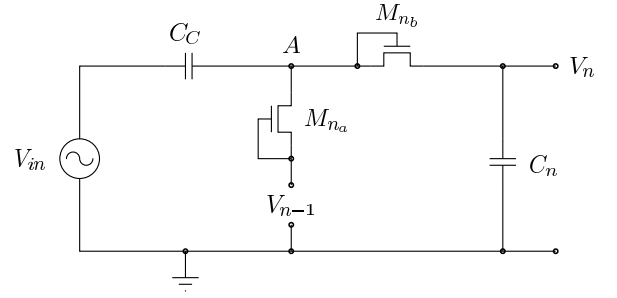


Fig. 5. One step of the charge pump

This charge pump works as follows; the positive half period of the input will charge capacitor C_n and the diode coupled transistor will prevent the current flowing back. At the zerocrossings of the input sine wave,

$$V_A = V_{n-1} - V_{th}. \quad (2)$$

At the peaks, this voltage rises $2 \cdot V_{in}$, and reaches

$$V_A = V_{n-1} - V_{th} + 2 \cdot V_{in} \quad (3)$$

C_n will be charged to $V_A - V_t$:

$$V_n = V_{n-1} + 2 \cdot V_{in} - 2 \cdot V_{th} \quad (4)$$

Each step of the charge pump can not increase the pumped voltage more than twice the amplitude of the input voltage minus twice the threshold voltage of the transistors. If the low-threshold models in the 90 nm technology are used, the diode threshold voltage will be 0.18 V. This means that, with no modifications of the circuit, the input amplitude must exceed 0.18 V to get the charge pump to pump voltage through the steps towards the last capacitor. With the backgate connection proposed in section III, the charge pump can work at even lower inputs.

To maximize the output voltage in this circuit, there are two main factors; minimize the threshold voltage or increase the number of steps. Since increasing the number of steps will cause degradation of power and conversion efficiency and also increased circuit size, lowering the voltage drop of the transistors is the most efficient way to improve the charge pump.

If we take a look at the transistor current in the sub-threshold region [5], we can see the relationship between the transistor current and W/L.

$$I_{ds} = I_{D0} \cdot \frac{W}{L} \cdot e^{(qV_{GS}/nkT)} \quad (5)$$

$$n = \frac{C_{ox} + C_{depl}}{C_{ox}} = 1.5 \quad (6)$$

To ensure pumping at low input voltages we will need to ensure that we are able to deliver as much current as possible. W/L and capacitors must be as large as possible without the capacitive load at the input of the circuit becoming too large. This can potentially be a problem in a practical

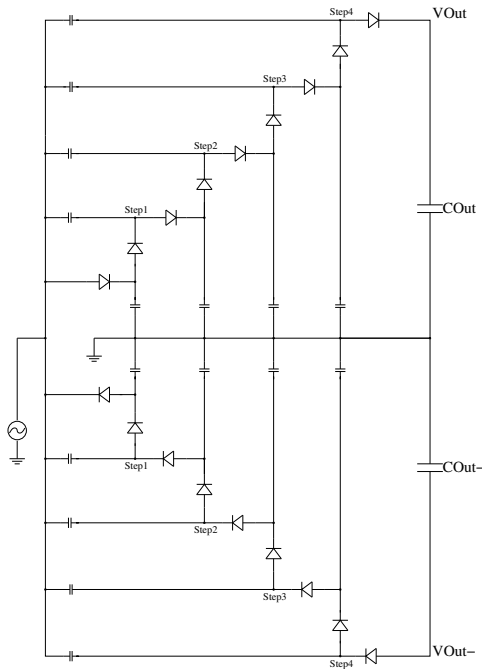


Fig. 6. Charge pump with both a positive and a negative output voltage

implementation, because the input source will then not be able to drive the charge pump.

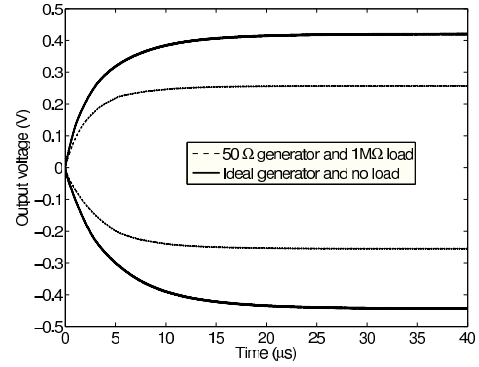
Until now only a positive output voltage has been extracted from the charge pump, but by further developing the pump like in Fig. 6, it can give a negative voltage of the same magnitude [6]. By defining V_{Out-} as reference, the available output voltage will be doubled. By doing this, the charge pump circuit will see a substrate potential which is equal to V_{Out-} and not its gnd node. This can cause some changed behaviour, and will have to be taken into consideration in a practical scenario.

V. CMOS CHARGE PUMP SIMULATIONS

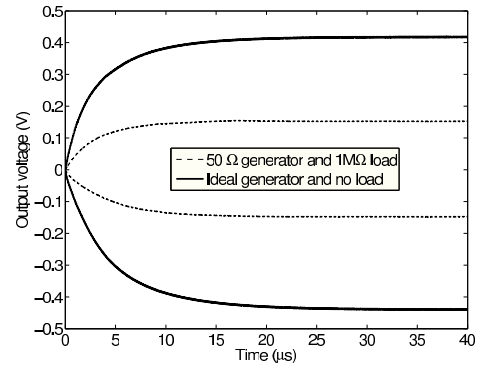
The charge pump is simulated with 900 MHz and 2.45 GHz input, and the results are shown in Fig. 7(a) and Fig. 7(b). These two frequencies are used by GSM and WLAN, and in this range several unlicensed bands can be found [7], [8]. The amplitude of the input signal is 100 mV. At higher frequencies the load will affect the circuit more and more, and will have to be kept in mind in a practical scenario. The results are promising for use in low-power wireless systems that require low voltage DC supplies.

VI. IMPLEMENTATION IN 90 NM CMOS

In Fig. 8 the layout of the chargepump is shown. It is arranged with the C_C capacitors on the left and the C_n capacitors on the right. The diodes are placed in the middle to avoid long wires. Two test outputs are placed in the middle of the pump for measuring purposes.



(a) 900 MHz



(b) 2.45 GHz

Fig. 7. Time development of charge pump output

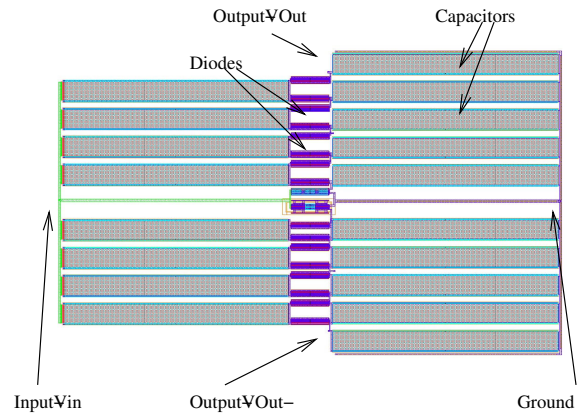


Fig. 8. Layout of the charge pump

VII. CHARGE PUMP MEASUREMENT SETUP AND PERFORMANCE

In Fig. 9 the PCB is shown. The long production time for the process used (>6 months) has delayed our measurements. The results so far are good, and matches the simulations. In Fig. 10(a) and 10(b) the output as a function of input signal is shown. In Fig. 10(c) the output power at different input amplitudes and output voltages are shown. As we can see in Fig. 10(a), the symmetry is broken when V_- is below approximately 0.5 V. This is because the pad frame of the chip has diodes between the pads and ground to prevent

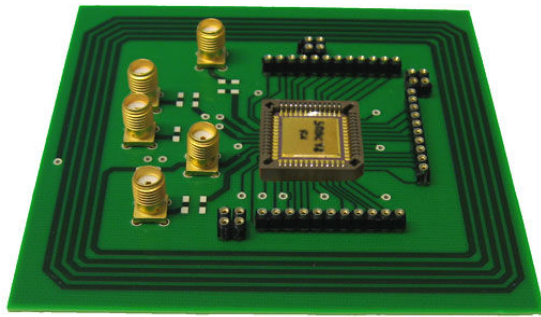


Fig. 9. The PCB with the chip mounted.

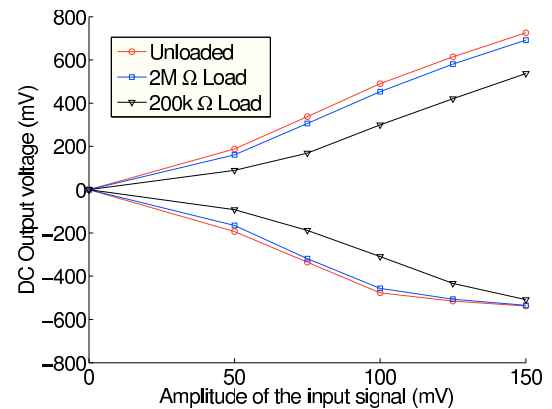
overloading of the chip. This problem can be solved by using a modified pad. As Fig. 10(c) shows, the charge pump is very power efficient. I.e. the pump require approximately 300 mV input signal to deliver 60 μ W at 1 V DC output. This can be sufficient to power a pulse transmitter, as a wireless identification tag.

VIII. CONCLUSION

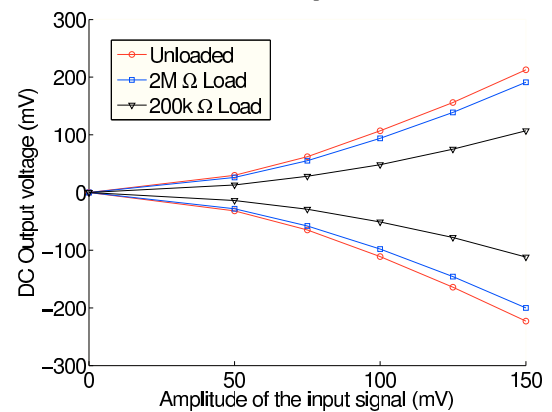
In this paper we have presented a high sensitivity charge pump suitable for power harvesting of reasonably small RF signals commonly emitted by wireless equipment. The novel exploration of the back-gate for improved performance indicates the feasibility of power harvesting in devices like RFID. In spite of presumed leakage in fine-pitch MOS devices, we are able to explore back-gate functions for power efficient charge pumps in 90nm technology.

REFERENCES

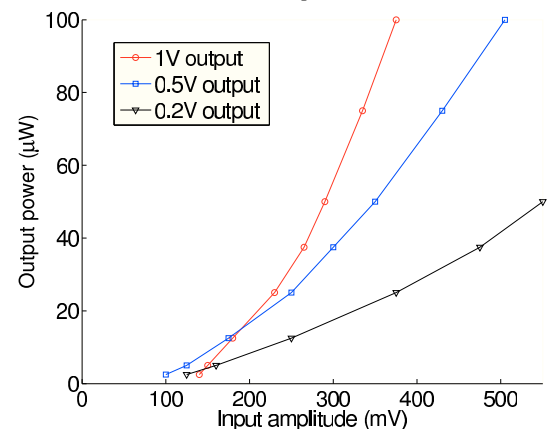
- [1] K. Finkenzeller, *RFID Handbook: Fundamentals and Applications in Contactless Smart Cards and Identification*, 2nd ed. John Wiley and Sons, 2003.
- [2] W. M. Sansen, *Analog Design Essentials*, 1st ed. Springer, 2006.
- [3] J. F. Dickson, "On-chip high-voltage generation in nmos integrated circuits using an improved voltage multiplier technique," *IEEE J. Solid-State Circuits*, vol. 11, no. 6, pp. 374–378, June 1976.
- [4] U. Karthaus and M. Fischer, "Fully integrated passive uhf rfid transponder ic with 16.7- μ W minimum rf input power," *IEEE J. Solid-State Circuits*, vol. 38, no. 10, pp. 1602–1608, October 2003.
- [5] D. Johns and K. Martin, *Analog Integrated Circuit Design*. John Wiley & Sons, Inc, 1997.
- [6] J.-P. Curty, N. Joehl, C. Dehollain, and M. J. Declercq, "Remotely powered addressable uhf rfid integrated system," *IEEE J. Solid-State Circuits*, vol. 40, no. 11, pp. 2193–2202, November 2005.
- [7] H. Leeson, P. Hansell, J. Burns, , and Z. Spasojevi, "Demand for use of the 2.4ghz ism band final report," Spectrum Management Advisory Group, Aegis Spectrum Engineering, Tech. Rep., July 2000.
- [8] *ERC Report 25 - European Common Allocation Table*. European Radiocommunications Committee (ERC), May 2004.



(a) Measured DC output at 1 GHz



(b) Measured DC output at 2.45 GHz



(c) Measured output power as a function of input and output voltages.

Fig. 10. Results of the measurements.

Acronyms

AA	Analog Average
A/D	Analog-to-Digital
ADC	Analog-to-Digital Converter
ATA	AT Attachment
AT	Advanced Technology
AWGN	Additive White Gaussian Noise
BER	Bit Error Rate
BPSK	Binary Phase-Shift Keying
BW	BandWidth
CD	Compact Disc
CDF	Cumulative Distribution Function
CDMA	Code Division Multiple Access
CMOS	Complementary Metal-Oxide Semiconductor
CPU	Central Processing Unit
CTBV	Continuous-Time Binary-Value
DAC	Digital-to-Analog Converter
D/A	Digital-to-Analog
DC	Direct Current
DFF	D Flip-Flop
DFT	Discrete Fourier Transform
DOA	Direction Of Arrival
DOI	Digital Object Identifier

DSD	Direct Stream Digital
DSSS	Direct-Sequence Spread Spectrum
EM	ElectroMagnetic
FCC	Federal Communications Commission
FFT	Fast Fourier Transform
GPS	Global Positioning System
HDR	High-Dynamic-Range
HDTV	High-Definition TeleVision
IF	Intermediate Frequency
IoT	Internet of Things
IR-UWB	Impulse-Radio Ultra-WideBand
ISAR	Inverse Synthetic Aperture Radar
LED	Light-Emitting Diode
LNA	Low Noise Amplifier
LSB	Least Significant Bit
M2M	Machine to Machine
MLE	Maximum-Likelihood Estimation
MOSFET	Metal-Oxide-Semiconductor Field-Effect Transistor
MSB	Most Significant Bit
MSE	Mean Squared Error
NMOS	N-channel MOSFET
PCA	Polarity Coincidence Array
PCIe	Peripheral Component Interconnect Express
PCM	Pulse-Code Modulation
PDF	Probability Density Function
PDM	Pulse-Density Modulation
PFM	Pulse-Frequency Modulation

PMF	Probability Mass Function
PMOS	P-channel MOSFET
PPM	Pulse-Position Modulation
PSD	Power Spectral Density
PVT	Process, Voltage, and Temperature
PWM	Pulse-Width Modulation
QAM	Quadrature Amplitude Modulation
RALA	RAce Logic Architecture
RC	Radio-Controlled
RC	Resistor-Capacitor
RF	Radio Frequency
RMSE	Root-Mean-Square Error
RMS	Root Mean Square
SACD	Super Audio CD
SATA	Serial ATA
SDM	Sigma-Delta Modulation
SDTV	Standard-Definition TeleVision
SINAD	Signal-to-Noise-And-Distortion ratio
SNR	Signal-to-Noise Ratio
SPAD	Single-Photon Avalanche Diode
SPI	Serial Peripheral Interface
SQNR	Signal-to-Quantization-Noise Ratio
SQUID	Superconducting QUantum Interference Device
SR	Stochastic Resonance
SSR	Suprathreshold Stochastic Resonance
ST	Swept Threshold
USB	Universal Serial Bus

UWB	Ultra-WideBand
WHO	World Health Organization
WSN	Wireless Sensor Network
XNOR	eXclusive Negated OR

Bibliography

- [Akyi 02] I. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci. “A survey on sensor networks”. *IEEE Communications Magazine*, Vol. 40, No. 8, pp. 102–114, Aug. 2002. doi:10.1109/MCOM.2002.1024422. Cited on page 18.
- [Aram 05] T. Arampatzis, J. Lygeros, and S. Manesis. “A Survey of Applications of Wireless Sensors and Wireless Sensor Networks”. In: *Proceedings of the 2005 IEEE International Symposium on, Mediterrean Conference on Control and Automation Intelligent Control, 2005*, pp. 719–724, Limassol, June 2005. doi:10.1109/.2005.1467103. Cited on page 18.
- [Barb 10] Y. Barbotin and P. C. Budianu. “Emulation of N-Bits Uniform Quantizer from Multiple Incoherent and Noisy One-Bit Measurements”. Patent Application, Sep. 2010. US 2010/0239055 A1. Available from: <http://www.google.com/patents?id=g9zWAAAAEBAJ>. Cited on page 79.
- [Beau 86] N. C. Beaulieu. *Bit error rate performance analysis and optimization of suboptimum detection procedures*. PhD thesis, University of British Columbia, 1986. Available from: <http://hdl.handle.net/2429/26775>. Cited on pages 112 and 135.
- [Benz 81] R. Benzi, A. Sutera, and A. Vulpiani. “The mechanism of stochastic resonance”. *Journal of Physics A: Mathematical and General*, Vol. 14, No. 11, pp. L453–L457, Nov. 1981. doi:10.1088/0305-4470/14/11/006. Cited on page 68.
- [Benz 82] R. Benzi, G. Parisi, A. Sutera, and A. Vulpiani. “Stochastic resonance in climatic change”. *Tellus*, Vol. 34, No. 1, pp. 10–16, 1982. doi:10.1111/j.2153-3490.1982.tb01787.x. Cited on page 68.
- [Braa 99] M. Braasch and A. J. Van Dierendonck. “GPS receiver architectures and measurements”. *Proceedings of the IEEE*, Vol. 87, No. 1, pp. 48–64, Jan. 1999. doi:10.1109/5.736341. Cited on pages 99 and 135.
- [Case 93] N. M. Casey and J. A. Angus. “One-Bit Digital Processing of Audio Signals”. In: *Audio Engineering Society Convention 95*, Audio Engineering Society, Oct. 1993. Available from: <http://www.aes.org/e-lib/browse.cfm?elib=6515>. Cited on pages 81 and 134.

- [DeVr 97] S. H. DeVries and D. A. Baylor. “Mosaic Arrangement of Ganglion Cell Receptive Fields in Rabbit Retina”. *Journal of Neurophysiology*, Vol. 78, No. 4, pp. 2048–2060, Oct. 1997. Available from: <http://jn.physiology.org/content/78/4/2048>. Cited on page 82.
- [East 97] P. C. Eastty, C. Sleight, and P. D. Thorpe. “Research on Cascadable Filtering, Equalization, Gain Control, and Mixing of 1-Bit Signals for Professional Audio Applications”. In: *Audio Engineering Society Convention 102*, Audio Engineering Society, March 1997. Available from: <http://www.aes.org/e-lib/browse.cfm?elib=7335>. Cited on page 134.
- [Fern 10] J. Fernandes and D. Wentzloff. “Recent advances in IR-UWB transceivers: An overview”. In: *Proceedings of 2010 IEEE International Symposium on Circuits and Systems (ISCAS)*, pp. 3284–3287, Paris, May 2010. doi:10.1109/ISCAS.2010.5537916. Cited on page 22.
- [Fuji 99] H. Fujisaka, N. Masuda, M. Sakamoto, and M. Morisue. “Arithmetic circuits for single-bit digital signal processing”. In: *The 6th IEEE International Conference on Electronics, Circuits and Systems, 1999. Proceedings of ICECS '99*, pp. 1389–1392, Pafos, Cyprus, Sep. 1999. doi:10.1109/ICECS.1999.850722. Cited on page 134.
- [Hami 14] T. Hamilton, S. Afshar, A. van Schaik, and J. Tapson. “Stochastic Electronics: A Neuro-Inspired Design Paradigm for Integrated Circuits”. *Proceedings of the IEEE*, Vol. 102, No. 5, pp. 843–859, May 2014. doi:10.1109/JPROC.2014.2310713. Cited on page 69.
- [Hari 09] V. N. Hari, G. V. Anand, and A. B. Premkumar. “Optimal suprathreshold stochastic resonance based nonlinear detector”. In: *17th European Signal Processing Conference (EUSIPCO 2009)*, pp. 2062–2066, Glasgow, Scotland, Aug. 2009. Available from: <http://www.eurasip.org/Proceedings/Eusipco/Eusipco2009/contents/papers/1569190766.pdf>. Cited on page 99.
- [Hjor 06] H. A. Hjortland. *UWB impulse radar in 90 nm CMOS*. Master’s thesis, Univ. of Oslo, 2006. Available from: <http://urn.nb.no/URN:NBN:no-12777>. Cited on pages 15, 78, and 138.
- [Hjor 07] H. A. Hjortland, D. T. Wisland, T. S. Lande, C. Limbodal, and K. Meisal. “Thresholded samplers for UWB impulse radar”. In: *Circuits and Systems, 2007. ISCAS 2007. IEEE International Symposium on*, pp. 1210–1213, New Orleans, LA, May 2007. doi:10.1109/ISCAS.2007.378326. Cited on page 78.
- [Hjor 09] H. A. Hjortland and T. S. Lande. “CTBV Integrated Impulse Radio Design for Biomedical Applications”. *Biomedical Circuits and Systems, IEEE Transactions on*, Vol. 3, No. 2, pp. 79–88, Apr. 2009. doi:10.1109/TBCAS.2009.2014962. Cited on pages 15 and 138.

- [Kane 66] M. Kanefsky. “Detection of weak signals with polarity coincidence arrays”. *IEEE Transactions on Information Theory*, Vol. 12, No. 2, pp. 260–268, Apr. 1966. doi:10.1109/TIT.1966.1053870. Cited on pages 29, 112, 134, and 135.
- [Kosk 01] B. Kosko and S. Mitaim. “Robust stochastic resonance: Signal detection and adaptation in impulsive noise”. *Physical Review E*, Vol. 64, No. 5, p. 051110, Oct. 2001. doi:10.1103/PhysRevE.64.051110. Cited on page 99.
- [Kosk 04] B. Kosko and S. Mitaim. “Robust stochastic resonance for simple threshold neurons”. *Physical Review E*, Vol. 70, No. 3, p. 031911, Sep. 2004. doi:10.1103/PhysRevE.70.031911. Cited on page 99.
- [Lee 02] S.-J. Lee and H.-J. Yoo. “Race logic architecture (RALA): a novel logic concept using the race scheme of input variables”. *IEEE Journal of Solid-State Circuits*, Vol. 37, No. 2, pp. 191–201, Feb. 2002. doi:10.1109/4.982425. Cited on page 141.
- [Max 60] J. Max. “Quantizing for minimum distortion”. *IRE Transactions on Information Theory*, Vol. 6, No. 1, pp. 7–12, March 1960. doi:10.1109/TIT.1960.1057548. Cited on page 90.
- [Maye 02] G.-F. Mayer-Lindenberg. “Programmable 1-bit data processing arrangement”. Patent Application, May 2002. US6385717 B1. Available from: <http://www.google.com/patents/US6385717>. Cited on page 66.
- [McDo 05] M. D. McDonnell, N. G. Stocks, C. E. M. Pearce, and D. Abbott. “Analog-to-digital conversion using suprathreshold stochastic resonance”. *Proceedings of SPIE*, Vol. 5649, No. 1, pp. 75–84, Feb. 2005. doi:10.1117/12.582493. Cited on page 110.
- [McDo 06a] M. D. McDonnell, N. G. Stocks, C. E. Pearce, and D. Abbott. “Optimal information transmission in nonlinear arrays through suprathreshold stochastic resonance”. *Physics Letters A*, Vol. 352, No. 3, pp. 183–189, March 2006. doi:10.1016/j.physleta.2005.11.068. Cited on page 78.
- [McDo 06b] M. D. McDonnell. *Theoretical Aspects of Stochastic Signal Quantisation and Suprathreshold Stochastic Resonance*. PhD thesis, The University of Adelaide, Australia, Feb. 2006. Available from: <http://www.eleceng.adelaide.edu.au/Personal/mmcdonne/Thesis/index.html>. Cited on pages 68, 78, 84, 93, and 110.
- [McDo 09a] M. McDonnell and N. Stocks. “Suprathreshold stochastic resonance”. *Scholarpedia*, Vol. 4, No. 6, p. 6508, 2009. doi:10.4249/scholarpedia.6508. Cited on page 69.
- [McDo 09b] M. D. McDonnell and D. Abbott. “What Is Stochastic Resonance? Definitions, Misconceptions, Debates, and Its Relevance to Biology”. *PLoS Comput Biol*, Vol. 5, No. 5, p. e1000348, May 2009. doi:10.1371/journal.pcbi.1000348. Cited on pages 67 and 69.

- [Niel 07] J. Nielsen. “UWB Impulse Radio Receiver Based on Single Bit Quantization”. *Wireless Personal Communications*, Vol. 43, No. 2, pp. 369–388, Oct. 2007. doi : 10.1007/s11277-006-9229-0. Cited on pages 134 and 135.
- [Papa 98] H. Papadopoulos, G. Wornell, and A. Oppenheim. “Low-complexity digital encoding strategies for wireless sensor networks”. In: *Proceedings of the 1998 IEEE International Conference on Acoustics, Speech and Signal Processing, 1998*, pp. 3273–3276, Seattle, WA, May 1998. doi : 10.1109/ICASSP.1998.679563. Cited on pages 78 and 79.
- [Prip 03] A. A. Priplata, J. B. Niemi, J. D. Harry, L. A. Lipsitz, and J. J. Collins. “Vibrating insoles and balance control in elderly people”. *The Lancet*, Vol. 362, No. 9390, pp. 1123–1124, Oct. 2003. doi : 10.1016/S0140-6736(03)14470-4. Cited on page 69.
- [Reml 66] W. R. Remley. “Some Effects of Clipping in Array Processing”. *The Journal of the Acoustical Society of America*, Vol. 39, No. 4, pp. 702–707, 1966. doi : 10.1121/1.1909944. Cited on pages 99, 112, and 122.
- [Rouv 07] C. Rouvas-Nicolis and G. Nicolis. “Stochastic resonance”. *Scholarpedia*, Vol. 2, No. 11, p. 1474, 2007. doi : 10.4249/scholarpedia.1474. Cited on pages 68 and 69.
- [Roy 06] V. M. Roy and G. V. Anand. “A suprathreshold stochastic resonance based pre-processor for direction-of-arrival estimation”. In: *TENCON 2006. 2006 IEEE Region 10 Conference*, pp. 1–4, Hong Kong, Nov. 2006. doi : 10.1109/TENCON.2006.343688. Cited on pages 99 and 111.
- [Sbai 09] L. Sbaiz, F. Yang, E. Charbon, S. Susstrunk, and M. Vetterli. “The gigavision camera”. In: *IEEE International Conference on Acoustics, Speech and Signal Processing, 2009. ICASSP 2009*, pp. 1093–1096, Taipei, Apr. 2009. doi : 10.1109/ICASSP.2009.4959778. Cited on page 135.
- [Sche 08] B. Schell and Y. Tsvividis. “A Continuous-Time ADC/DSP/DAC System With No Clock and With Activity-Dependent Power Dissipation”. *IEEE Journal of Solid-State Circuits*, Vol. 43, No. 11, pp. 2472–2481, Nov. 2008. doi : 10.1109/JSSC.2008.2005456. Cited on page 138.
- [Stan 08] J. Stankovic. “Wireless Sensor Networks”. *Computer*, Vol. 41, No. 10, pp. 92–95, Oct. 2008. doi : 10.1109/MC.2008.441. Cited on page 18.
- [Stew 98] R. Stewart and E. Pfann. “Oversampling and sigma-delta strategies for data conversion”. *Electronics & Communication Engineering Journal*, Vol. 10, No. 1, pp. 37–47, Feb. 1998. doi : 10.1049/ecej : 19980106. Cited on page 81.
- [Vuor 13] T. Vuori, J. Alakarhu, E. Salmelin, and A. Partinen. “Nokia PureView oversampling technology”. In: *Proc. SPIE*, pp. 86671C_1–86671C_10, SPIE-IS&T, Burlingame, California, USA, Feb. 2013. doi : 10.1117/12.2020064. Cited on page 82.

- [Widr 96] B. Widrow, I. Kollar, and M.-C. Liu. “Statistical theory of quantization”. *IEEE Transactions on Instrumentation and Measurement*, Vol. 45, No. 2, pp. 353–361, Apr. 1996. doi:10.1109/19.492748. Cited on pages 78 and 128.
- [Wiki 12] Wikipedia. “Stochastic resonance (sensory neurobiology) — Wikipedia, The Free Encyclopedia”. 2012. [Online; accessed 02-October-2012]. Available from: [http://en.wikipedia.org/w/index.php?title=Stochastic_resonance_\(sensory_neurobiology\)&oldid=510532575](http://en.wikipedia.org/w/index.php?title=Stochastic_resonance_(sensory_neurobiology)&oldid=510532575). Cited on page 69.
- [Wiki 14] Wikipedia. “Oversampled binary image sensor — Wikipedia, The Free Encyclopedia”. 2014. [Online; accessed 07-August-2014]. Available from: http://en.wikipedia.org/w/index.php?title=Oversampled_binary_image_sensor&oldid=610389700. Cited on page 135.
- [Wiki 15] Wikipedia. “Johnson-Nyquist noise — Wikipedia, The Free Encyclopedia”. 2015. [Online; accessed 31-May-2015]. Available from: http://en.wikipedia.org/w/index.php?title=Johnson%E2%80%93Nyquist_noise&oldid=654488506. Cited on page 45.
- [Wu 15] T. Wu, T. Rappaport, and C. Collins. “Safe for Generations to Come: Considerations of Safety for Millimeter Waves in Wireless Communications”. *IEEE Microwave Magazine*, Vol. 16, No. 2, pp. 65–84, March 2015. doi:10.1109/MMM.2014.2377587. Cited on page 23.
- [Zeng 13] X. Zeng, S. Ni, R. Ge, and F. Wang. “Design of Nonuniform Quantizer in Satellite Navigation Receivers”. In: W. Lu, G. Cai, W. Liu, and W. Xing, Eds., *Proceedings of the 2012 International Conference on Information Technology and Software Engineering*, pp. 119–128, Springer Berlin Heidelberg, Jan. 2013. Available from: http://dx.doi.org/10.1007/978-3-642-34522-7_14. Cited on pages 99 and 135.
- [Zier 12] C. Zierhofer. “Signal Representation With Unity-Weight Dirac Impulses”. *IEEE Transactions on Signal Processing*, Vol. 60, No. 6, pp. 2860–2869, June 2012. doi:10.1109/TSP.2012.2190729. Cited on page 140.